Volume 6, No. 2; April 2018

Advances in Image And Video Processing

ISSN: 2054-7412

Society for Science and Education - United Kingdom

TABLE OF CONTENTS

EDITORIAL ADVISORY BOARD

DISCLAIMER

All the contributions are published in good faith and intentions to promote and encourage research activities around the globe. The contributions are property of their respective authors/owners and the journal is not responsible for any content that hurts someone's views or feelings etc.

AIVP ADVANCES IN IMAGE AND VIDEO PROCESSING
SOCIETY FOR SCIENCE AND EDUCATION UNITED KINGDOM
ISSN 2054-7412

Spectral and Spatial Quality Assessment of IHS and Wavelet Based Pan-sharpening Techniques for High Resolution Satellite Imagery

Farzaneh DadrasJavan, Farhad Samadzadegan, Fatemeh Fathollahi
Department of Geomatics, University College of Engineering, University of Tehran, Tehran, Iran
fdadrasjavan@ut.ac.ir, samadz@ut.ac.ir, fz.fathollahi@ut.ac.ir

ABSTRACT

Over last decades, a wide range of pan-sharpening methods have been proposed to synthesize images in a way that contain both high spatial and spectral characteristics of the multispectral and panchromatic images in high resolution satellite imagery. Amongst all improved methodologies, two different scenarios of Wavelet based and IHS based strategies brought up to be more appealing. Until now, lots of modifications and integrations are also proposed for these methods and variety of new techniques are presented. In this paper, the potential of different IHS-based and wavelet-based pan-sharpening techniques is studied and evaluated. For the purpose, high resolution images of worldview-2 satellite imagery are used and pan-sharpened images are generated based on 9 wavelet based and 6 IHS based methods. Achieved results clearly show the superiority of GIHS method as an optimum solution to keep both spectral and spatial characteristics if input images in fused image.

Keywords: Satellite imagery, Image Fusion, Spatial and Spectral quality, comparative study, HIS, Wavelet

1 Introduction

The application of satellite images is still growing due to the rapid development in sensor technologies. Technological limitations in producing images with both high spectral and spatial resolutions, led to the pan-sharpening (i.e. image fusion) process which produces synthesized multispectral image with both high spectral and spatial resolution. Considering the importance of pan-sharpening process, till now variety of pan-sharpening methods have been proposed and applied to various high resolution satellite imagery (Yuhendra et al., 2012; Asha & Philip, 2012; Amolins et al., 2007, Zhang and Huang, 2015; Liu et al., 2017).

One of the common ideas in pan-sharpen image generation, is to extract the spatial details from the Pan image and to add them somehow to the MS bands, those suffer from lack of spatial information. Multi resolution analysis is known to be the most common tool for this purpose. Variety of image fusion procedures have been proposed using the multi resolution analysis based on the discrete wavelet transform (DWT). Wavelet transformation is employed to decompose images into different resolution levels for multi resolution analysis which can be utilized for fusing images at different resolution levels (Zhang & Hong, 2005, Amolins et al., 2007).

On the other hand, there are also other widely used techniques which are mainly named as component substitution methods (Dou et al., 2007). The main procedure of these methods is to apply a predefined transformation to multispectral bands and then to substitute one of the obtained components by Pan

DOI: 10.14738/aivp.62.4226
Publication Date: 24th ,February 2018
URL: http://dx.doi.org/10.14738/aivp.62.4226

image. Component substitution methods usually yield to images with high spatial qualities while the wavelet methods usually perform well spectrally (Kim et al., 2011). Based on these differences, many researches try to combine these two procedures (Kim et al., 2011; González-Audícana et al., 2005; Zhang & Hong, 2005). The most proposed schemes combine the advantages of wavelets and Intensity-Hue-Saturation (IHS) or wavelets and Principal Component Analysis (PCA) based fusion methods. The idea to merge the procedures of wavelets with HIS or PCA is to improve the results spectrally by injecting the high frequency detail information of Pan to the component I/PC1 instead of replacing Pan with I or PC1 component of the spectral image (I is the intensity of multispectral image and PC1 is the first principal component of Pan image).

In this paper, a short review on wavelet and IHS based fusion methods is presented and using worldview-2 satellite imagery, their capabilities in generating pan-sharpen images with high spatial and spectral resolutions, are investigated. Moreover, both spatial and spectral quality metrics are developed to compare fusion results.

2 Review on Pan-sharpening Techniques and Evaluation Metrics

Despite the diversity in current pan-sharpening methods, most of them can generally be classified into Component Substitution (CS) and Multi-Resolution Analysis (MRA) methods (Zhang and Huang, 2015). The Component Substitution based techniques commonly used in literatures are: Intensity Hue Saturation, Principal Component Analysis (PCA), Arithmetic Combinations and Brovey Method. These techniques are simple and are reported to provide better spatial quality in final fused image (Kolekar and Shelkikar, 2016). Amongst these methods, colour transform based fusion methods (such as HIS) have become popular mainly because of their simplicity, fast computational process and also ability to preserve spatial characteristic of the Pan image (Tu et al., 2004; Choi, 2006). On the other hand, MRA techniques involve transformation of source images into different scales before applying fusion rules. The commonly used MRA techniques are pyramid, wavelet, Contourlet and Curvelet transforms (Kolekar and Shelkikar, 2016). A popular multi resolution transform is the Discrete Wavelet Transform (DWT) hich is used for pixel-level image fusion algorithms (Kolekar and Shelkikar, 2016).

2.1 Wavelet-Based Pan-Sharpening Techniques

Numerous fusion methods are developed based on the discrete wavelet transform (DWT) those are usually different in the algorithm of image decomposition or in the way of detail information injection, i.e. substitution, addition or weighted models (Amolins et al., 2007). The Undecimated Discrete Wavelet, known as “a` trous” (meaning “with holes”), yields to shift-invariant discrete wavelet decomposition which is required in the application of image fusion Kim et. al., 2011; Mitchell 2010; Amolins et. al., 2007).

The process of pan-sharpening based on wavelets, consists of three main stages. Firstly, the input images should be transformed to the wavelet domain. After transformation, detail coefficients (related to high frequency information of the image in different scales) and an approximation image (related to low frequency information of the image at the last scale) will be obtained. Then in the second step, the detail and approximation coefficients of the decomposed images are combined to synthesize the fused decomposition coefficients. The third step is to reconstruct the final fused image using the inverse wavelet transformation (Mitchell, 2010; Amolins et al., 2007, Kolekar and Shelkikar, 2016). In each step, some important decisions such as the scheme of wavelet decomposition and the method to combine the coefficients should be made before the fusion process.

2.1.1 Decomposition Stage

There are several different methods to apply the wavelet transformation or in the other word to decompose an image into wavelet coefficients. The Mallat's algorithm uses an orthonormal basis with decimated approach, but the transformation is not shift-invariant, which can be a problem in data fusion. Instead, "a` trous" scheme which is an undecimated discrete wavelet transformation, yields to shift-invariant discrete wavelet decomposition (Amolins et al., 2007; González-Audícana et al., 2005; Rockinger, 1996). In this paper the undecimated wavelet which is a non-orthogonal, redundant transformation is applied.

2.1.2 Methods to Combine Different Coefficients

The important stage in the wavelet-based fusion process is how to combine the coefficients of the decomposed Pan and MS images or how to inject the detail information to MS bands. The most common methods are substitution, addition or weighted models (Amolins et al., 2007).

- Substitution. In substitution schemes, after applying wavelet transformation, the detail coefficients (or wavelet planes) of one image are substituted by the detail coefficients of the other image coefficients (Amolins et al., 2007; Kim, et. al., 2011).
- Addition. When applying undecimated wavelet transformation, the input images have the same size as the wavelet planes. So one of the input images can be directly added to the wavelet planes obtained after applying undecimated wavelet transformation to the second image (Amolins et al., 2007).
- Weighted Models. In IHS method, a weighted model is used to combine either the approximation or the detail coefficients obtained after applying wavelet transformation to both input images (González-Audícana et al., 2004).

2.1.3 Different Pan-sharpening Schemes

Most well-known wavelet-based pan-sharpening methods with the 'a` trous' decomposition scheme can be summarized as follows:

- Substitutive Wavelet (SW), (Amolins et al., 2007; Kim et al., 2011).
- Additive Wavelet (AW) (Amolins et al., 2007).
- Substitute Wavelet Intensity (SWI) (González-Audícana et al., 2004).
- Additive Wavelet Intensity (AWI) (Nunez et al., 1999).
- Additive Wavelet Luminance Proportional (AWLP) (Otazu et al., 2005; Kim et al., 2011).
- Weighted Wavelet Intensity (WWI) (Zhang & Hong, 2005).
- Substitute Wavelet Principal Component (SWPC) (González-Audícana et al., 2004).
- Additive Wavelet Principal Component (AWPC) (González-Audícana et al., 2005).

In Table 1 at jth level, W_{Pan_j} is the wavelet plane of the Pan. I in Table 1 refers to Intensity component obtained after the IHS transformation while PC1 refers to first principal component obtained after principal component transformation. Also LL refers to the approximation image obtained at the last level of wavelet decomposition, P refers to Pan and I refers to Intensity image and w1 and w2 are calculated according to equation 1.

$$LL' = w1 \times LL^P + w2 \times LL^I$$
$$w1 = corr(LL^P, LL^I)\ , \qquad w1 + w2 = 1 \tag{1}$$

2.2 IHS-Based Pan-Sharpening Techniques

The IHS method has a standard procedure while its major limitation is in keeping spectral content of the MS bands, which is famous as color/spectral distortion. Also it has limitation of using only three bands at once. It has been proved that the main reason for spectral distortion is the difference between the spectral response of the Pan image and intensity component of MS image (Tu et al., 2001). Different solutions have been proposed to overcome spectral distortion in color transform based methods by minimizing the Pan and Intensity difference and also to modify the method in order to be capable of using an arbitrary number of spectral bands (Tu et al., 2007; Choi, 2006, Aiazzi and Baronti, 2007, Xu et al., 2008).

These techniques can be divided into three main groups. First group, taking into account the spectral response plot of the sensor, have extended traditional three dimensional IHS to use all the bands whose spectral range falls within the range of the Pan image (Tu et al., 2004). To improve IHS extension, some coefficients have been proposed for each MS bands just based on try and error (Tu et al., 2004; Choi, 2006). Second group researches have tried to determine the quota of each MS band using the area under the spectral response functions (González-Audícana et al., 2006) and the third group have applied the mathematical methods to approximate Pan by means of the available MS bands (Aiazzi et al., 2007; Xu et al., 2008; Rahmani et al., 2010).

In IHS based fusion methods, after applying IHS transformation on the image, intensity (I), hue (H) and saturation (S) will be obtained with the I component to appear like Pan image (Zhang & Hong, 2005). Using IHS characteristic, pan-sharpened MS bands can be achieved through the following steps (Tu et al., 2001; Zhang & Hong, 2005):

a) Apply IHS transformation to RGB set of resampled MS image.
b) Perform histogram matching between Pan and I.
c) Replace I component by modified Pan image.
d) Apply inverse IHS transformation to data
e) The simplified and more efficient form of the above stages which avoids numerous multiplications is as follows (Tu et al., 2004):

$$F_i = MS_i + \delta \tag{2}$$

Where $\delta = \mathrm{Pan} - \mathrm{I}$ and $\mathrm{F_i}$ and MS_i refer to the ith band of the fused and multispectral images respectively. Tu et al. (2001) proved that the reason of the color distortion is the change of δ component during fusion process. He also demonstrated that the difference between I and Pan leads to the change of δ. To IHS problem, the literature shows some suggestions in which the main effort is to modify I to be more similar to Pan image. IHS methods mostly applied for fusion are summarized as follow Equations are presented in Table 2):

- Generalized IHS (GIHS) (Tu et al., 2004, 2007): considering the relative spectral response of the IKONOS and the fact that the spectral range of the Pan covers also Near Infra-Red (NIR) band besides R, G and B bands have extended the traditional three-order transformation to four-order one.
- Generalized IHS with Tradeoff Parameter (GIHS-TP), Choi, (2006) has proposed the minimization problem for the IHS method and introduced I_{new} .
- GIHS with Best Tradeoff Parameter (GIHS-BTP), Tu et al., (2007) has modified Choi method using a simple energy-normalization procedure. Tu et al. demonstrated that their approach results to better preservation of spatial information than that of Choi's method.

- GIHS with Adaptive Weights (GIHS-AW), Aiazzi and Baronti (2007) have proposed to perform a linear regression between Pan and MS bands. In the proposed scheme, a synthetic intensity, having a minimum mean square error (MSE) with respect to the reduced Pan (i.e. at the spatial scale of the original MS image), is computed. The intensity I have been assumed as a linear combination of MS bands (equation 6) with unknown coefficients (w_i) and a bias (b) which are computed using a linear regression algorithm (Aiazzi et al., 2007).
- Improved GIHS with Adaptive Weights (IGIHS-AW), the method can extend traditional three-order transformations to an arbitrary order and all the MS bands could be fused at the same time (Xu et al., 2008).

2.3 Spectral Quality Assessment Metrics

Spectral image fusion quality assessment metrics are classified based on the level of spectral information, considered in quality assessment process (Alparone et al., 2006; Thomas and Wald, 2006). Traditionally, these metrics are classified as mono-modal and multi-modal techniques (Xydeas and Petrovic, 2000). A mono-modal metric considers a single image layer while a multi-modal metric considers several image channels.

Mono modals: Correlation Coefficients (CC) between the original and the fused images; Universal Quality Index (UQI) which measures the salient information contained in reference image which has been transformed into the fused image (Thomas and Wald, 2006);

Multi Modals: Spectral Angle Mapper (SAM) that obtains the angles formed (α) between the reference spectrum and the fused image spectrum (Leung et al., 2001); The ERGAS index or Relative Adimensional Global Error in the fusion (Wald, 2000; Riahi et al., 2009); and Spectral Correlation Mapper (SCM) (Van der meer, 2006).

The quality assessment metrics are presented in Table 3 where x and y stand for fused and reference images, $\bar{x}$ and $\bar{y}$ are the local sample means of x and y, σx and σy are the local sample standard deviations of x and y, and σxy is the sample cross correlation of x and y after removing their means. S is the number of spectral bands of images or the dimension of the spectral space, A = (A1, A2, A3,…, AS) and B=(B1, B2, B3,…, BS) are two spectral vectors with same wavelength from the reference and fused images respectively. Mi is the mean value for the original images, Bi represents bands of images and GSDPan and GSDMS stand for the ground sampling resolution of initial panchromatic and multi spectral images, respectively

2.4 Spatial Quality Assessment Metrics

In addition to preserving the spectral information of the input MS image, a fused image should contain the spatial information of the Pan image, as well, especially in high resolution satellite imagery which are usually used for urban area mapping (Javan et al., 2013). The spatial metrics usually are used for pan-sharpening quality evaluation are presented as CC, UQI, spatial ERGAS (ERGASs) proposed by Lillo-Saavedra et al. (2005) and spatial CC (sCC) proposed by Zhoue et al. (1998).

3 Experiments and Results

To evaluate and compare the capabilities of discussed image fusion methods, all are applied for Pan-sharpening of WV2 satellite imagery. The evaluation strategy is presented in Figure 1. As it's depicted, in the first step, initial Pan and MS images are down sampled and Low resolution Pan (L Pan) and multi spectral (L MS) images are generated. Using this new images, Pansharpen image (Fused L image) is generated based on Wald strategy (). Then quality assessment is applied comparing the generated fused image with initial multi spectral image (H MS) spectrally and with initial Pan image (H Pan) spatially.

3.1 Data Set

The WV2 data set which is from an urban area of Melbourne city (Figure 2). WV2 has 8 spectral bands of 500×500 pixels and a Pan band of 2000×2000 pixels with 2 m and 0.5 m resolutions, respectively. The MS bands are up-sampled and the pairs of images were geometrically registered to each other. A part of images (500×500 pixels) is extracted to compare the visual effects of different fused images.

3.2 Wavelet Based Image Fusion Results

In IHS experiment, the undecimated transformation scheme was used to decompose the input images in all the methods and the level of decomposition was set to 2 for all the resultant images. To compare the visual effect of the different fused images, a part of the resultant images (500×500 pixels) is extracted and shown in Figure 3. From Figure 3, it can be seen that the spatial resolutions of all the resultant images are improved through the fusion process in comparison with the original MS. Considering the resultant images of additive schemes, it seems that they suffer from some distortions along the lines in the images exactly similar to those that can be seen along the edges of the original MS. The reason is that in additive schemes, the low resolution image (MS itself or Intensity component or PC1) is directly added to the detail coefficients of the Pan without being decomposed.

3.3 IHS based Image Fusion Results

Considering the spectral response function of the WV2 sensors (depicted in Figure 4) IHS-based methods for WV2 can profit from the additional yellow and red edge bands whose spectral range fall within the range of the Pan (Figure4). So, in the modified scenario for WV2, the five bands of blue, green, yellow, red and red edge were used in the process to improve the spectral quality of the fused image. In IHS experiment all the mentioned methods reviewed in chapter 2, were applied to WV2 data. To compare the effects of the fusion process visually, a part of resultant images (500×500 pixels) is extracted and shown in Figure 5. From results, although visual comparison of images is not convenient, but it is clearly observed that all methods improved the spatial quality of MS image. But some of the methods like GIHSAW and IHS resulted in better spatial quality while IHS presented spectral distortion.

3.4 Spectral Quality Assessment

To evaluate the spectral quality of the fused images, the Pan and original MS images were degraded to a resolution of 2 and 4 m respectively and were fused at a 4m resolution. The spectral quality was evaluated by comparing the spectral information of fused image to that of the original MS image applying spectral quality assessment metrics. Table 4 and 5 show the results of spectral metrics calculated to evaluate the fused images.

To provide the capability of visual comparison of results, they are also depicted in Figure 6. To avoid messy figure, results are presented for the mean value of R, G and B bands. Comparing the values in Table 4 and Figure 6, it can be concluded that the method SWPC, while averaging all its bands, yields to a fused image with superior spectral quality. However, it can be seen that there is no significant difference between results of SWPC and SWI methods, which shows the success of these hybrid methods in preservation of spectral information. It is also observed that substitution schemes usually result to better preservation of spectral content of the original MS image in comparison with additive schemes (see Figures 6.b and 6.c).

From Table 5, it can be concluded that all mentioned IHS based methods have better spectral quality in comparison with traditional IHS method and the superior spectral quality was obtained by IGIHS-AW

method. As IHS method is designed to minimize spectral distortion by keeping the spectral vector of the fused image parallel to that of the original MS, the obtained value for SAM is near to zero.

3.5 Spatial Quality Assessment

Figure 7 presents the results of pan-sharpening methods based on spatial metrics while comparing generated pan-sharpen and Pan images. The values of ERGASs, UQI and sCC calculated for three spectral bands (R, G and B) are averaged and demonstrated in Figure 7.

From Tables 6 and 7 also Figure 7, all calculated metrics, address the WWI method as the superior method in terms of spatial quality while the sCC metric has addressed the SW method. It shows that the idea to keep the Pan approximation partially, instead of ignoring it all, can improve the spatial quality of the fused image. Comparing the additive schemes with substitutive ones, it can be found that the metrics sCC and UQI show better spatial qualities by the results of additive scheme, while the metrics sCC and ERGASs, show the better functionality of the substitutive schemes in keeping the spatial information. It seems that the metrics sCC and ERGASs are more sensitive to those distortion effects of additive schemes along the lines that can be easily observed in the results.

4 Discussions

6 common IHS based fusion methods and 9 different wavelet based fusion methods are evaluated based on spectral and spatial quality assessment metrics. Results clearly show the total spatial dominance of IHS based methods and spectral superiority of Wavelet based Methods. Results also revealed that in wavelet based method SWPC and WWI show the best results for spectral and spatial quality respectively. Moreover, among IHS based methods, IGIHS-AW and traditional IHS methods present the best for spectral and spatial quality respectively.

On the other hand, to have better sense over the results which are discussed in previous section, we applied an averaging strategy of evaluation and comparison. For the purpose, the values for each metric are normalized based on their global minimum and maximum values in each table of 4, 5, 6 and 7. Then, the normalized values are averaged for each method in both case of spectral and spatial assessment as a grade. Results are presented in table 8 and 9. The best quality results are marked in both spectral and spatial case. Moreover, as in lots of applications both spectral and spatial qualities are important, users should apply methods fairly enjoy both spectral and spatial characteristics. So, the grades in each case are again normalized with respect to their min/max values and averaged and the final score for each of these wavelet based and IHS methods are computed.

As it is presented in tables 8 and 9, while considering spectral quality, GIHS and for spatial quality, IHS presents the best results. Moreover if both spectral and spatial qualities are important, GIHS method can perform better. Although it is not the best spectrally or spatially, it can somehow present both in an acceptable level.

5 Conclusion

The widespread use of pan-sharpening in generation of high spatial and spectral quality in high resolution satellite imagery, has led to a wide range of image fusion techniques. In this paper the potential of wavelet based and IHS based techniques as most common Pan-sharpening methods have been inspected. 6 IHS based and 9 Wavelet based methods were developed and their capabilities in generating fused images were studied based on Worldview-2 satellite imagery. Obtained results clearly showed that if spectral resolution is considered, SWPC and GIHS are the most wisdom selection. On the other hand, in case of spatial resolution the classic IHS and WWI methods are the best. Moreover, as both spectral and spatial quality are important in lots of application, results showed that GIHS and

SW is one of the best which showed acceptable results in both cases and can be applied for lots of Pan-sharpening applications.

REFERENCES

[1] Aiazzi, B., S. Baronti, and M. Selva. "Improving Component Substitution Pansharpening Through Multivariate Regression of MS+ Pan Data." IEEE Transactions on Geoscience and Remote Sensing 45 (2007): 3230-3239.

[2] Alimuddin, I., Sumantyo, J. T. S., & Kuze, H. (2012). Assessment of pan-sharpening methods applied to image fusion of remotely sensed multi-band data. International Journal of Applied Earth Observation and Geoinformation, 18, 165-175.

[3] Amolins, K., Zhang, Y., & Dare, P. (2007). Wavelet based image fusion techniques—An introduction, review and comparison. ISPRS Journal of Photogrammetry and Remote Sensing, 62(4), 249-263.

[4] Asha, G., and Annes Philip. "A review on pixel level satellite image fusion." Int. J. Comput. Appl 1 (2012): 15-19.

[5] Choi, M. (2006). A new intensity-hue-saturation fusion approach to image fusion with a tradeoff parameter. IEEE Transactions on Geoscience and Remote sensing, 44(6), 1672-1682.

[6] Javan, F. D., Samadzadegan, F., & Reinartz, P. (2013). Spatial quality assessment of pan-sharpened high resolution satellite imagery based on an automatically estimated edge based metric. Remote Sensing, 5(12), 6539-6559.

[7] Dou, W., Chen, Y., Li, X., & Sui, D. Z. (2007). A general framework for component substitution image fusion: An implementation using the fast image fusion method. Computers & Geosciences, 33(2), 219-228.

[8] González-Audícana, M., Otazu, X., Fors, O., & Alvarez-Mozos, J. (2006). A low computational-cost method to fuse IKONOS images using the spectral response function of its sensors. IEEE Transactions on Geoscience and Remote Sensing, 44(6), 1683-1691.

[9] Kim, Y., Lee, C., Han, D., Kim, Y., & Kim, Y. (2011). Improved additive-wavelet image fusion. IEEE Geoscience and Remote Sensing Letters, 8(2), 263-267.

[10] Kolekar, M. N. B., & Shelkikar, R. P. A . (2016). A Review on Wavelet transform based image fusion and classification. International Journal of Application or Innovation in Engineering & Management (IJAIEM), 5 (3), 111-115

[11] Leung, L. W., King, B., & Vohora, V. (2001, November). Comparison of image data fusion techniques using entropy and INI. In Paper presented at the 22nd Asian Conference on Remote Sensing (Vol. 5, p. 9).

[12] Liu, X., Wang, Y., & Liu, Q. (2017). Remote Sensing Image Fusion Based on Two-stream Fusion Network. arXiv preprint arXiv:1711.02549.

[13] Lillo-Saavedra, M., & Gonzalo, C. (2006). Spectral or spatial quality for fused satellite imagery? A trade-off solution using the wavelet à trous algorithm. International Journal of Remote Sensing, 27(7), 1453-1464.

[14] Mallat, S. G. (1989). A theory for multiresolution signal decomposition: the wavelet representation. IEEE transactions on pattern analysis and machine intelligence, 11(7), 674-693.

[15] Mitchell, H. B. (2010). Image fusion: theories, techniques and applications. Springer Science & Business Media.

[16] Nunez, J., Otazu, X., Fors, O., Prades, A., Pala, V., & Arbiol, R. (1999). Multiresolution-based image fusion with additive wavelet decomposition. IEEE Transactions on Geoscience and Remote sensing, 37(3), 1204-1211.

[17] Rahmani, S., Strait, M., Merkurjev, D., Moeller, M., & Wittman, T. (2010). An adaptive IHS pan-sharpening method. IEEE Geoscience and Remote Sensing Letters, 7(4), 746-750.

[18] Riyahi, R., Kleinn, C., & Fuchs, H. (2009). Comparison of different image fusion techniques for individual tree crown identification using QuickBird images. International Society for Photogrammetry and Remote Sensing, High-Resolution Earth Imaging for Geospatial Information, 38, 1-4.

[19] Rockinger, O., & Fechner, T. (1998, April). Pixel-level image fusion: the case of image sequences. In Proc. SPIE (Vol. 3374, pp. 378-388).

[20] Thomas, C., & Wald, L. (2006, July). Analysis of changes in quality assessment with scale. In Information Fusion, 2006 9th International Conference on (pp. 1-5). IEEE.

[21] Tu, T. M., Huang, P. S., Hung, C. L., & Chang, C. P. (2004). A fast intensity-hue-saturation fusion technique with spectral adjustment for IKONOS imagery. IEEE Geoscience and Remote sensing letters, 1(4), 309-312.

[22] van der Meer, F. (2006). The effectiveness of spectral similarity measures for the analysis of hyperspectral imagery. International journal of applied earth observation and geoinformation, 8(1), 3-17.

[23] Wald, L. (2000, January). Quality of high resolution synthesised images: Is there a simple criterion?. In Third conference" Fusion of Earth data: merging point measurements, raster maps and remotely sensed images"(pp. 99-103). SEE/URISCA.

[24] Xu, J., Guan, Z., & Liu, J. (2008). An improved IHS fusion method for merging multi-spectral and panchromatic images considering sensor spectral response. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci, 37, 1169-1174.

[25] Zhang, Y., & Hong, G. (2005). An IHS and wavelet integrated approach to improve pan-sharpening visual quality of natural colour IKONOS and QuickBird images. Information Fusion, 6(3), 225-234.

[26] Zhang, H. K., Huang, B., Zhang, M., Cao, K., & Yu, L. (2015). A generalization of spatial and temporal fusion methods for remotely sensed surface parameters. International Journal of Remote Sensing, 36(17), 4411-4445.

[27] Zhang, H. K., & Huang, B. (2015). A new look at image fusion methods from a Bayesian perspective. Remote Sensing, 7(6), 6828-6861.

[28] Zhou, J., Civco, D. L., & Silander, J. A. (1998). A wavelet transform method to merge Landsat TM and SPOT panchromatic data. International Journal of Remote Sensing, 19(4), 743-757.

Table 1: Specification of wavelet-based methods

Method	Detail Coefficients	Approximation Coefficients
SW	$\sum_{j=1}^{3} W_{Pan_j}$	LLMS
AW	$\sum_{j=1}^{3} W_{Pan_j}$	MS
SWI	$\sum_{j=1}^{3} W_{Pan_j}$	LLI
AWI	$\sum_{j=1}^{3} W_{Pan_j}$	I
SWPC	$\sum_{j=1}^{3} W_{Pan_j}$	LLPC1
AWPC	$\sum_{j=1}^{3} W_{Pan_j}$	PC1
AWLP	$\frac{MS_i}{\frac{1}{3}\sum_{i=1}^{3} MS_i} \sum_{j=1}^{3} W_{Pan}$	I
WWI	$\sum_{j=1}^{3} W_{Pan_j}$	$w1 \times LL^{Pan} + w2 \times LL^{I}$

Table 2: Specification of IHS based methods

Method	Equation
Generalized IHS (GIHS)	$(R + G + B + NIR)/4$
Generalized IHS with Tradeoff Parameter (GIHS-TP)	$I_{new} = Pan - \frac{Pan - I}{t}$ $F_i = MS_i + \delta_{choi}$ $\delta_{choi} = I_{new} - I = \frac{t-1}{t}(Pan - I)$
GIHS with Best Tradeoff Parameter (GIHS-BTP)	$F_i = \frac{Pan}{I_{new}}(MS_i + \delta_{choi})$
GIHS with Adaptive Weights (GIHS-AW)	$I = \sum_{i=1}^{n} w_i . MS_i + b$
Improved GIHS with Adaptive Weights (IGIHS-AW)	$I = \sum_{i=1}^{n} w_i . MS_i$
Improved GIHS with Adaptive Weights (IGIHS-AW)	$F_i = MS_i + \frac{MS_i}{I}\delta$

Table 3: Fusion quality assessment metrics

Correlation Coefficients	$corr = \frac{\sum_{1}^{M}\sum_{1}^{N}(x-\bar{x}).(y-\bar{y})}{\sqrt{\sum_{1}^{M}\sum_{1}^{N}(x-\bar{x})^2.\sum_{1}^{M}\sum_{1}^{N}(y-\bar{y})^2}}$
Universal Quality Index	$Q = \frac{\sigma_{xy}}{\sigma_x.\sigma_y}.\frac{2.\bar{x}\bar{y}}{\bar{x}^2+\bar{y}^2}.\frac{2\sigma_x.\sigma_y}{\sigma_x^2+\sigma_y^2}$
Spectral Angle Mapper (SAM)	$cos(\alpha) = \frac{\sum_{i=1}^{N} A_i.B_i}{\sqrt{\sum_{i=1}^{N} A_i.A_j}.\sqrt{\sum_{i=1}^{N} B_i.B_j}}$
ERGAS	$ERGAS = 100\frac{GSD_{Pan}}{GSD_{MS}}\sqrt{\frac{1}{S}\sum_{i=1}^{S}\left[\frac{RMSE(B_i)^2}{\bar{x}_i^2}\right]}$
Spectral Correlation Mapper	$C_{xy} = \frac{\sum_{1}^{N}(x-\bar{x}).(y-\bar{y})}{\sqrt{\sum(x-\bar{x})^2.\sum(y-\bar{y})^2}}$

Table 4: Spectral quality metrics calculated for different Wavelet based fusion methods

		SW	AW	SWI	AWI	SWPC	AWPC	AWRGB	AWLP	WWI
CC	R	0.97	0.96	0.97	0.97	0.974	0.970	0.94	0.97	0.96
	G	0.97	0.96	0.97	0.96	0.974	0.970	0.94	0.96	0.96
	B	0.97	0.96	0.97	0.96	0.974	0.970	0.95	0.96	0.96
UQI	R	0.68	0.68	0.72	0.71	0.697	0.694	0.60	0.71	0.68
	G	0.72	0.71	0.72	0.71	0.727	0.720	0.59	0.70	0.71
	B	0.63	0.63	0.62	0.62	0.655	0.651	0.60	0.60	0.63
ERGAS		5.04	9.07	6.15	4.91	6.09	4.85	5.88	8.92	5.37
SAM		5.33	13.8	6.41	4.53	6.347	4.37	6.69	13.48	5.46
SCM		0.92	0.88	0.91	0.93	0.922	0.93	0.93	0.87	0.91

Table 5: Spectral quality metrics calculated for different IHS based fusion methods

		IHS	GIHS	GIHS-TP	GIHS-BTP	GIHS-AW	IGIHS-AW
CC	R	0.86	0.94	0.96	0.94	0.96	0.97
	G	0.84	0.92	0.95	0.93	0.95	0.97
	B	0.80	0.92	0.95	0.92	0.95	0.97
UQI	R	0.55	0.65	0.72	0.66	0.66	0.79
	G	0.53	0.64	0.72	0.65	0.64	0.80
	B	0.42	0.53	0.61	0.55	0.55	0.80
ERGAS		12.0	7.95	6.36	7.69	6.37	5.28
SAM		16.9	5.00	3.96	3.96	4.25	0.02
SCM		0.58	0.79	0.86	0.80	0.90	0.93

Table 6: Spatial quality metrics calculated for different Wavelet based fusion methods.

Metrics / Methods		SW	AW	SWI	AWI	SWPC	AWPC	AWRGB	AWLP	WWI
C	R	0.993	0.935	0.933	0.868	0.868	0.885	0.922	0.89	0.902
	G	0.941	0.943	0.933	0.932	0.933	0.944	0.931	0.95	0.958
	B	0.882	0.885	0.921	0.920	0.922	0.934	0.869	0.82	0.946
UQI	R	0.811	0.799	0.638	0.635	0.635	0.683	0.762	0.77	0.743
	G	0.807	0.804	0.767	0.762	0.765	0.803	0.761	0.84	0.852
	B	0.693	0.685	0.751	0.750	0.757	0.797	0.641	0.67	0.834
sCC	R	0.993	0.969	0.974	0.971	0.974	0.965	0.977	0.995	0.991
	G	0.994	0.970	0.981	0.978	0.981	0.971	0.978	0.981	0.997
	B	0.989	0.966	0.978	0.976	0.978	0.970	0.973	0.972	0.994
ERGASs		11.4	11.94	11.55	11.9	11.56	11.9	12.9	13	10.18

Table 7: Spatial quality metrics calculated for different IHS based fusion methods.

Method / Metric		IHS	GIHS	GIHS-TP	GIHS-BTP	GIHS-AW	IGIHS-AW
CC	R	0.97	0.94	0.92	0.94	0.85	0.85
	G	0.99	0.98	0.97	0.98	0.92	0.92
	B	0.97	0.89	0.87	0.88	0.77	0.77
UQI	R	0.86	0.86	0.82	0.86	0.73	0.69
	G	0.85	0.86	0.88	0.86	0.82	0.78
	B	0.75	0.75	0.71	0.74	0.63	0.57
sCC	R	0.99	0.98	0.98	0.98	0.98	0.95
	G	0.99	0.99	0.98	0.99	0.98	0.96
	B	0.99	0.99	0.99	0.99	0.96	0.91
ERGASs		4.6	7.5	8.4	7.7	12.0	11.9

Table 8: Summarized grades of Wavelet based methods

	SW	AW	SWI	AWI	SWPC	AWPC	AWRGB	AWLP	WWI
Spectral Quality	0.61	0.63	0.66	0.56	0.73	0.64	0.21	0.63	0.49
Spatial Quality	0.68	0.44	0.47	0.41	0.43	0.44	0.47	0.55	0.71
Quality	0.83	0.45	0.53	0.33	0.53	0.46	0.10	0.63	0.76

Table 9: Summarized grades of IHS based methods

	IHS	GIHS	GIHS-TP	GIHS-BTP	GIHS-AW	IGIHS-AW
Spectral Quality	0.45	0.79	0.67	0.51	0.66	0.65
Spatial Quality	0.82	0.79	0.75	0.79	0.59	0.42
Quality	0.50	0.96	0.73	0.55	0.52	0.29

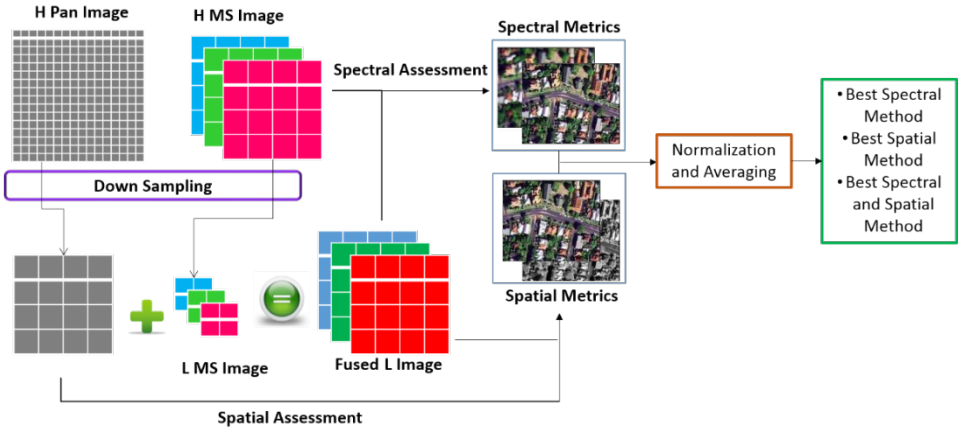
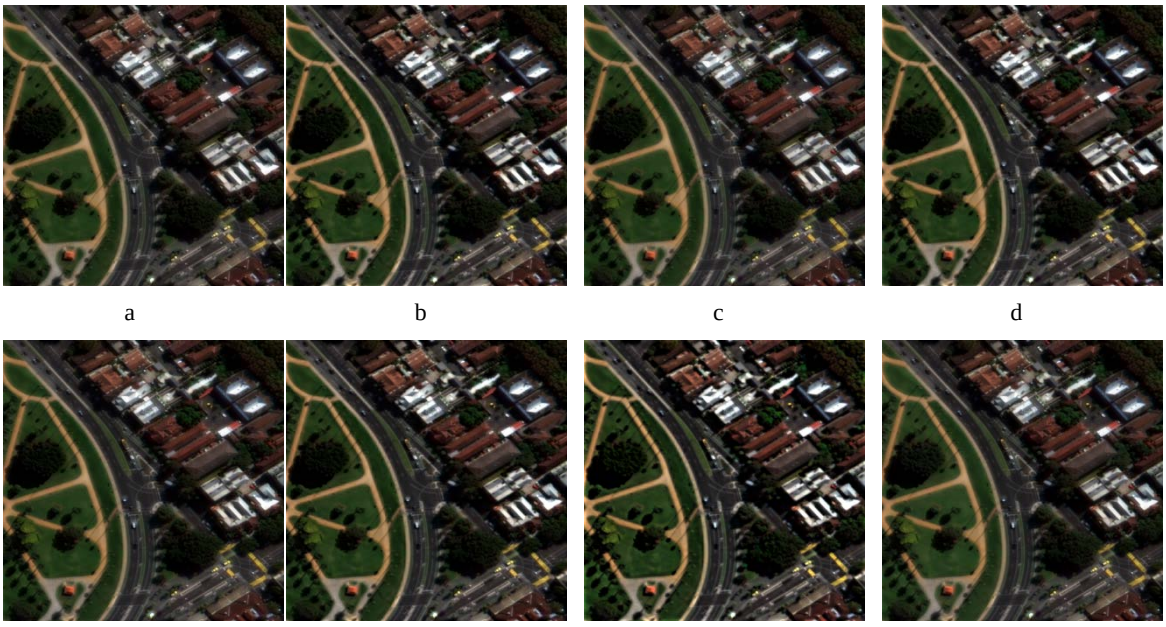


Figure 11: Image Fusion quality assessment strategy

a b c d

Figure 2: Image dataset. a) Original pan image with 0.5 m resolution. b) Selected part of original pan image (500×500 pixels). c) RGB combination of the up-sampled MS image with 2 m resolution. d) Selected part of the RGB combination of up-sampled MS (500×500 pixels).



a b c d

e f g h

Figure 3: Results of the applied wavelet-based pan-sharpening methods. a) SW b) AW c) SWI d) AWI e) SWPC f) AWPC g) AWLP h)WWI

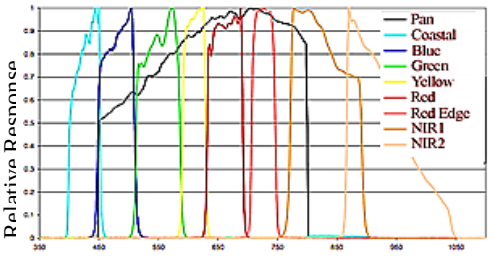
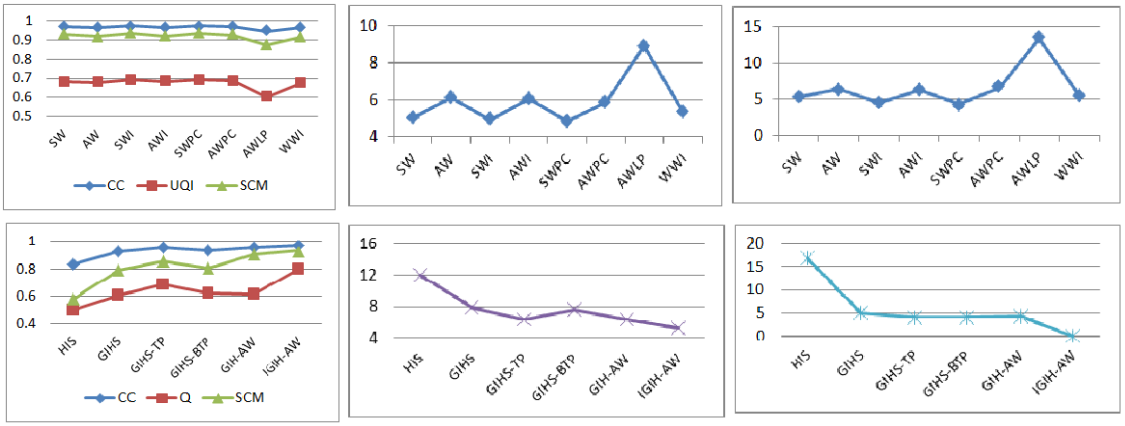


Figure 4: Spectral response of the WV2 sensors

a b c d

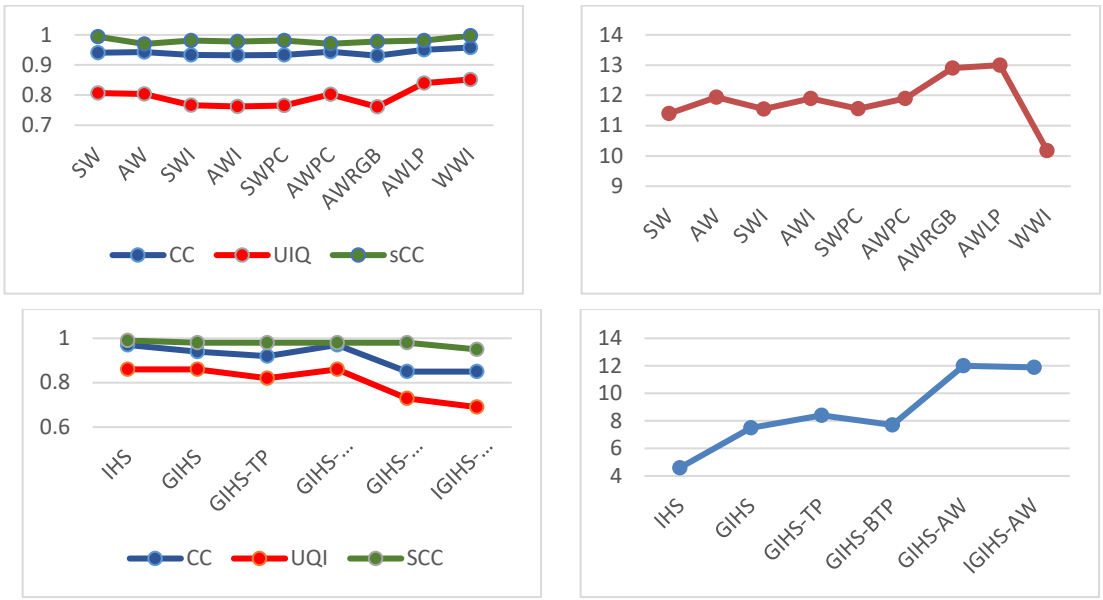
e f

Figure 5: Results of the IHS-based pan-sharpening methods. a) Traditional IHS b) GIHS c) GIHS-TP d) GIHS-BTP e) GIHSAW f) IGIHS-AW



a. CC, Q, SCM **b. ERGAS** **c. SAM**

Figure 6: Spectral quality metrics calculated for different methods. above) Wavelet based methods, below) IHS based methods



a) CC,UQI and sCC(averaged over three bands of R, G and B)

b) ERGASs

Figure 7: Spatial quality metrics calculated for different methods. above) Wavelet based methods, below) IHS based methods

AIVP ADVANCES IN IMAGE AND VIDEO PROCESSING
SOCIETY FOR SCIENCE AND EDUCATION UNITED KINGDOM
ISSN 2054-7412

Evaluation of Suitability of Voice Reading of Al-Qur'an Verses Based on Tajwid Using Mel Frequency Cepstral Coefficients (MFCC) and Normalization of Dominant Weight (NDW)

[1]Heriyanto, [2]Sri Hartati, [3]Agfianto Eko Putra
[1]Faculty of Teknic Industry (FTI), University of Pembangunan Nasional "Veteran" University, (UPN) Tambak bayan Babarsari Yogyakarta Indonesia;
[2,3]Departement of Computer Science and Electronics, Faculty of Mathematics and Natural Science (FMIPA), Gadjah Mada University (UGM), Sekip Bulaksumur, Yogyakarta 55281 Indonesia;
Heriyanto.yanto@mail.ugm.ac.id; shartati@ugm.ac.id; agfi@ugm.ac.id

ABSTRACT

The recitation of the Qur'an has its own uniqueness, among others having a special rule in reading and pronunciation, which is called tajwid science. At the time of the Qur'an is recited, there are often mistakes due to the limitations of knowledge of Tajwid. Therefore, the availability of tools to facilitate in checking the appropriateness of recitation is very much needed by those who recite the Qur'an and face limitations in understanding the science of tajwid. Checking the Qur'an reading is a problem that must be solved according to the rules. So far, voice identification studies have problems with feature extraction, compatibility or suitability testing, and accuracy. The issue of feature extraction, suitability, and impermanence testing have been improved in this study, which consists of two stages. The first stage is the extraction of the sound character of the Qur'an reading and the second stage is the testing of the conformity of the Qur'anic recitation and accuracy. In the first stage feature extraction is handled using MFCC and Normalization of Dominant Weight (NDW). Characteristics of reading the Qur'an as reference table is taken from one reader of Al-Qur'an who has competence in the field of science tajwid, for sampling 5-7 people as a source for testing. The process of the second stage of conformity testing of Qur'an reading is done starting from filtering, sequential multiplication of reference table and Conformity Uniformity Pattern (CUP). The sample of reading conformity test is taken from 11 Qur'anic letters containing 8 reading laws and 886 records. The test is performed on the dominant frame, the number of cepstral coefficient and the number of frames. The reading conformance test provides an average accuracy of 91.37% on the nine dominant frames. The test for the number of cepstral coefficients in the c-23 can be an average of 96.65%, while the number of frames on the F-10 is the best average of 96.65%.

Keywords: voice; reading; Al-Qur'an; MFCC; suitability; feature extraction; reference table

1 Introduction

The Qur'anic recitation sounds has its own uniqueness, among others having a special rule in reading and pronunciation by Zarkasyi [1], called tajwid science. Tajwid science is the ordinance in rightly reciting the Qur'an. At the time of the Qur'an is recited, there are often mistakes due to the limitations of knowledge of Tajwid. Therefore, the availability of tools to facilitate in checking the appropriateness of reading sounds is very much needed by those who read the Qur'an and they have limitations in understanding the science of tajwid. Correct reading of the Qur'an is a problem that requires its own

DOI: 10.14738/aivp.62.4268
Publication Date: 20th March 2018
URL: http://dx.doi.org/10.14738/aivp.62.4268

challenge, tailored to the rules. Research identification of the solution is done by testing matching or suitability and accuracy of reading. The issue of characteristic extraction, suitability testing, and accuracy were corrected in this study.

2 Literature Review

Research on voice identification is required for Voice To Text (VTT) applications. VTT research in Bahasa Indonesia by Suyanto and Hartati [2] identifies speech signals into vocabularies that include phonemes, syllables and accuracy segmentation of 98.93%. A similar study was conducted by Suyanto and Putra [3] using Mel Frequency Ceptrum Coefficient (MFCC) and Hidden Markov Model (HMM), which resulted in the introduction of phoneme segmentation in Indonesian. Another research conducted by Cahyarini et al.[4], which identifies the pause speech between phonemes that have an accuracy of 80%. Sound research in Nigerian by Yang et al. [5] identifies sayings 1 through 9, which results in 90% accuracy. Similar research involving words like "cash", "bought" and "man" was performed by Bodruzzaman et al. [6], with an accuracy of 90%. Sound research to control the robot as done by Chen et al. [5], in the form of the command words "front", "back", "right" and "left" with a level of accuracy of 80%. Robot greeting command study was also conducted by Tomasouw and Irawan [7] which resulted in a level of accuracy of 97.1%.

So far, feature extraction of the most widely used speech signals is Mel-Frequency Ceptrum Coefficients (MFCC) in both speaker recognition and speech recognition by Davis and Mermelstein [8]. Feature extraction with MFCC according to Aibinu et al. [9] has a level of accuracy of 58%. The recognition accuracy 75 % for MFCC by Hidayat [10].

To date, voice identification research has constraints ranging from feature extraction, testing and accuracy. This research is intended to improve the level of accuracy by developing the concept of extraction of Qur'anic recitations and the development of concepts for testing the appropriateness of reading the Qur'an.

3 Research Method

The evaluation of Al-Qur'an reading conformity study uses MFCC characteristic extraction and the Dominant Weighted Normalization model (NWD) in order to produce the features available in the reference reading conformity table. The reference table is derived from the source reading of its features and tested the reading of the source with the test reader. Records of recorded sources are from one expert who has competence in the field of science tajwid, while sample testers are taken from 5-7 people who are not experts. Source and tester samples were taken by recording with murotal reading. The murotal recording is a single word or two word verses of the Qur'an as an example of Al-Fatehah's letter reading one word in the first verse "Bismillah hirrohman nirrohim", the recording taken is "bismillah", "hirrohman" and "nirrohim". Taking a recording of two-word readings eg "ghoirill maghdhu bi alaihim", then the recording taken is "ghoiril magdhu". The Qur'anic letters taken are not all but samples of 11 letters containing 8 reading laws and 886 recording files. Furthermore, recording the recitation of the Qur'an is done extraction feature and testing the suitability of reading. Figure 1 shows a model for evaluating the conformity of the Qur'an reading.

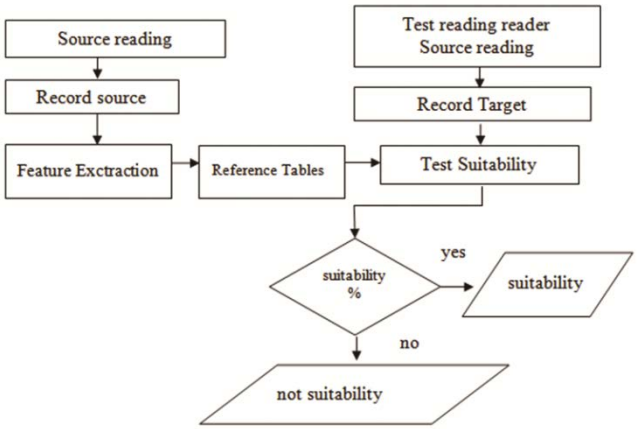


Figure 1:Model of conformity evaluation of Qu'anic reading

To conduct research the reading of Al-Qur'an readiness consists of two stages. The first stage feature extraction by using Mel Frequency Ceptrum Coefficient (MFCC) and developing the model of Normalization of Dominant Weight (NDW) to be able to produce a reference reading conformity table. In the second stage, a conformity test is performed by applying the Conformity Uniformity Pattern (CUP) to obtain a reading fit with a good degree of accuracy. Tests conducted are the dominant frame, the number of cepstral coefficient and the number of frames.

4 Feature Extraction

In the first stage feature extraction is performed to obtain the reference table using Mel Frequency Ceptrum Coefficients (MFCC) and the Dominant Weighted Normalization model (NDW). MFCC consists of 7 processes, namely preemphasing, framing blocking, windowing, Fast Fourier Transform (FFT), Mel Frequency Wrapping (MFW), Discrete Cosinus Transform (DCT) and Cepstral Liftering which produces cepstral coefficient feature. Then the cepstral coefficient is processed by using the Normalization of the Dominant Weight (NDW) to obtain the reference table for the suitability of the reading.

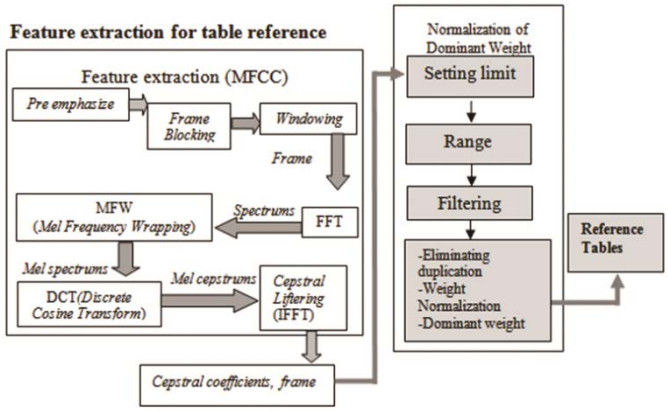


Figure 2:MFCC feature extraction section and NDW Model generate reference table

It is shown in Figure 2 that after going through the MFCC, then the NDW model is used in determining the limits, range, filtering, eliminating duplication, normalizing weights and dominant weights to generate reference tables. Here is an explanation of MFCC and NDW process stages.

4.1 Stages of MFCC

MFCC is a mef frequency and mel cepstral which produces the features of cepstral coefficients representing linear and non linear data in the form of cosine. Mel Frequency Cepstral Coefficents

(MFCCs) are a feature widely used in automatic speech and speaker recognition. They were introduced by Davis and Mermelstein in the 1980[11]. MFCC algorithm is used for voice recognition by Chakraborty [12]. Some of the advantages of this method according to Manunggal [13] is able to capture the voice characteristics that are very important for voice recognition, or in other words, it can capture important information contained in the voice signal. The MFCC stages are described as follows.

4.1.1 Preemphasis

The first stage in the MFCC process is preemphasis. Preemphasis is a noise signal filter that reduces noise, maintains a high signal. The preemphasize process is based on the equation y (n) = s (n) - α s (n - 1), with the value of α between 0 to 1. In general, the commonly used α values between 0.9 to 1.0 are based on following.

$$y(n) = s(n) - \alpha\, s(n-1) \tag{1}$$

Whereas, *1) y (n): signal result of preemphasize filter,*

2) s (n): signal before preemphasize filter,

3) n: order of samples,

4) α: coefficient of filter (0-1), s: signal.

4.1.2 Frame Blocking

Frame blocking is a process that divides the sound sample into multiple frames with a certain time and length. The time in the frame is taken between 10-30 milli seconds. Frame is taken as long as possible to get good frequency resolution, while shortest time may be intended to get the best time domain. The percentage of frame overlapping is between 30% - 50%. The frame blocking equation is as follows.

$$f(1; ns) = s^{|}(ns + ms(l-1)) \tag{2}$$

Whereas , *1) ns : 0..N-1,*

2) l : 1..L,

3) f : frame,

4) s : signal, dan ns, ms:sampel.

The standard timing takes 25 milliseconds with a sample rate of 44.100 Hz thus the sample size is equal to 25 milliseconds to 0.025 x 44.100 for 1,102.5 samples. The calculation also considers 50% overlap of 1,102.5 ie 551.25 samples so the frame rate is 44.100 / (1,102,5-551,25) = 80 frames per second

4.1.3 Windowing

Windowing is an interlocutor performed for each frame. Windowing aims to reduce the discontinuity effect at the ends of frames generated by the frame blocking process. Windowing consists of Triangular, Hamming and Hanning. The researchers apply windowing hanning because it is finer according to Agfianto [14]. The following is a representation of the window function of the input sound signal.

$$X(n) = Xi(n)w(n) \tag{3}$$

Whereas, *1) n : 0,1,…,N-1,*

2) X (n): sample signal value of windowing result,

3) Xi (n): the sample value of the frame signal to i,

4) w (n): window function,

5) N: frame size, is a multiple of 2.

The Hanning window function is as follows :

$$w(n) = 0.5(1-(\cos\frac{2x\pi i}{M-1})) \quad (4)$$

Whereas, *1) n : 0,1,..., M-1,*

2) M : frame length,

3) w(n) : function of window.

4.1.4 Fast Fourier Transform (FFT)

FFT is the process of signaling from time domain to frequency domain. Fourir is an enabling method for analysis of spectral properties that converts signals from time domain to frequency domain. FFT is a fast algorithm for implementing Discrete Fourier Transform (DFT). The DFT equation is as follows:

$$s[k] = \sum_{n=0}^{N-1} s(n) e^{-\frac{2\pi i}{N}nk} \quad 0 \le k \le N-1 \quad (5)$$

Whreas, *1)* N: *number of samples to be processed,*

2) s (n): the sample signal value,

3) k:The discrete frequency variable that is valuable (k=N/2,k ∈ N), n:0.1,…,M-1.

4.1.5 MFW (Mel Frequency Wrapping)

MFW is a mel scale of filterbank with frequency domain. Mel frequency is linear for frequencies below 1000 Hz and is logarithmic for frequencies above 1000 Hz. The purpose of MFW is to produce a mel spectrum. The following is the equation used in the calculation of filterbanks and MFW.

$$Y[i] = \sum_{j-1}^{Nm} s[j][Hi[j] \quad (6)$$

Whereas, *1) Nm: number of magnitude spectrum (N ∈ N),*

2) s [j]: magnitude spectrum at frequency j,

3) Hi [j]: coefficient filterbank at frequency j (1 ≤ i ≤ M),

4) M: the number of channels in the filterbank.

$Mel(f)$ = 2595 x log_{10} (1 + f/700) (7)

Whereas, *1)Mel* : *ScaleMel,*

2) f : frequency.

4.1.6 Discrete Cosine Transform (DCT)

DCT is a correlate of mel spectrum. DCT comes from FFT which in inverse get the value of liftering. The value of the liftering is processed again The second FFT yields the value of cepstrum. DCT goals produce mel cepstrum to improve recognition quality. The equation used to calculate DCT.

$$Cn = \sum (\log S_k) \cos\left[n(k-\frac{1}{2})\frac{\pi}{K} \right]; n = 1,2,...K \quad (8)$$

Whereas, *1) Cn: Coefficient of MFCC,*

2) Sk: output from filterbank process on index k,

3) K: number of expected coefficient.

4.1.7 Cepstral Liftering

Cepstral liftering is an implementation of the window function against cepstral coefficient features. The formula is as follows.

$$w(ni) = 1 + \frac{L}{2}\sin(\frac{nix\pi}{L}); ni = 1,2..L \quad (9)$$

Whereas, *1) w [ni]: window to cepstral features,*

2) L: number of cepstral coefficients,

3) ni: index of cepstral coefficients.

Generally, the feature of cepstral coefficients for one word taken as many as 11-12 cepstral coefficients Suyanto and Hartati [2]. Al-Qur'an reading study consisting of one or two words so that 2 x12 = 24 features cepstral coefficients.

Table 1. MFCC frame results and cepstral coefficients

Name	Reading	Frame	Coefficient to	Cepstral Coefficients
Barkoni-Iqra01.wav	Iqra	0	0	27.98
Barkoni-Iqra01.wav	Iqra	0	23	41.67
Barkoni-Iqra01.wav	Iqra	1	0	49.38
Barkoni-Iqra01.wav	Iqra	1	23	0.84
Barkoni-Iqra01.wav	Iqra	10	0	37.87
Barkoni-Iqra01.wav	Iqra	10	23	34.70

Table 1 is a feature extraction with MFCC in the form of cepstral coefficient and frame features. Cepstral coefficient is processed to the next stage by using the model of Normalization of Dominant Weight (NDW).

4.2 Stages of Normalization of Dominant Weight (NDW)

The Normalization Model of the Dominant Weight, hereinafter referred to as the NDW consisting of 6 stages, namely determining the limits, ranger, filtering, eliminating duplication, normalization of weight and dominant weight. The whole process is called the Normalization of Dominant Weight (NDW). The NDW model can be explained as follows.

4.2.1 Setting Limit

Limit is a line that becomes the boundary of a field, separator between two fields, one part with another part. Boundary selection is also based on the selection algorithm. The election algorithm is an algorithm to find the smallest number of k (the largest number of k) in a list. This algorithm is also called the order of statistics. The statistical order is used to determine the limits based on the features of cepstral coefficient starting from the smallest value, the largest value, the smallest value plus the largest value divided by two (median) and the mean value. The formulation is as follows.

$Limit1 = min$ *(initial limit),*
$Limit2 = (min + (min + max)/2)/2,$
$Limit3 = average,$

$$Limit4 = (min + max)/2,$$
$$Limit5 = (\left(\frac{min+max}{2}\right) + max)/2,$$
$$Limit6 = max \ (final\ limit). \quad (10)$$

The fetching limit becomes 6, as it uses the quarter and average system. Limit is divided into 5 parts consisting of 4 parts with quarter system and one part with average seen in Figure 3.

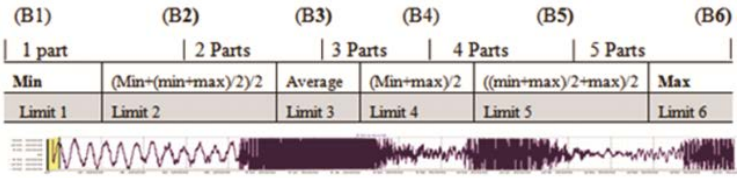


Figure 3:Illustration of limit 1 to limit 6

Limits are to be used as one of the features that exist in the reference table suitability reading. Figure 4 shows that the 1st limit is B1 up to the 6th or B6 boundary consisting of readings and frames.

MFCC→	Frame, cepstral coefficients		
Reading	frame	Coeffisients to	Cepstral Coefficients
Iqra	0	0	27,98
Iqra	0	23	41,67
Iqra	1	0	49,38
Iqra	1	23	0,84
Iqra	10	0	37,87
Iqra	10	23	34,7

1. Limits							
Reading	frame	B1	B2	B3	B4	B5	B6
Iqra	0	-35,13	2,67	22,33	40,48	78,30	116,11
Iqra	1	-18,05	13,5	28,24	45,24	76,88	108,53
Iqra	10	-25,12	6,16	24,97	37,46	68,75	100,04

Figure 4:Setting a limit of 1 to 6

4.2.2 Setting Range

Range is also called the span in the data variant. Range is the difference between the largest value data with the smallest value. Range is taken based on previously created limits consisting of 6 limits. The determination of the range is taken from the 1st or B1 limit up to the 6th or B6 limit resulting in several conditions formed with the following equation.

$$1.\ Condition\ of\ R1\ is\ if\ ((limit1)\ minimum\ value = cc)\ then\ P1Weight = 1,$$
$$2.\ Condition\ of\ R2\ is\ if\ (((limit1) \geq cc)\ and\ (cc < limit2))then\ P2Weight = 1,$$
$$3.\ Condition\ of\ R3\ is\ if\ (((limit2) \geq cc)\ and\ (cc < limit3))then\ P3Weight = 1,$$
$$4.\ Condition\ of\ R4\ is\ if\ (((limit3) \geq cc)\ and\ (cc < limit4))then\ P4Weight = 1,$$
$$5.\ Condition\ of\ R5\ is\ if\ (((limit4) \geq cc)\ and\ (cc < limit5))then\ P5Weight = 1,$$
$$6.\ Condition\ of\ R6\ is\ if\ (((limit5) \geq cc)\ and\ (cc < limit6))then\ P6Weight = 1,$$
$$7.\ Condition\ of\ R7\ is\ if\ \big((limit6) maximum\ value = cc\big) then\ P7Weight = 1\ \ . \quad (11)$$

In this case, *cc: features of cepstral coefficients.*

Based on the above equation then formed into 7 conditions based on range, which then used to the filtering process.

4.2.3 Filtering

Filtering is the process of filtering. Filtering is a feature of source data based on range. Filtering itself aims to separate the features of the data in order to collect on each part. The feature collection of the data in each section is named P1, P2, P3, P4, P5, P6 and P7. Filtering process that is done if it meets the

conditions according to range, then given weight 1 on each of P1 until P7. The subsequent process is added to each of P1 to P7 and the sum is made to the whole P1 to P7 called K.P. The equation is as follows.

$$K.P = \sum_{PJ1}^{PJ7}$$
$$PJ1 = \sum_{f0}^{fm} P1 \quad (12)$$

The equaton is also applicable to*PJ2, PJ3, PJ4, PJ5, PJ6, PJ7*

Whereas, *1) K.P: total number of PJ1 to PJ7,*

2) f0: the 0th frame,

3) fm: frame to-m,

4) PJ1: the number of P1,

5) PJ2: the number of P2,

6) PJ3: the number of P3,

7) PJ4: the number of P4,

8) PJ5: the number of P5,

9) PJ6: the number of P6,

10) PJ7: the number of P7,

11) P1, P2, P3, P4, P5, P6, P7: filtering results in the form of weight.

Figure 5 shows that filtering P1 through P7 with the 0th to-m frame taking turns there is duplication of one reading to another.

3.Filtering

Reading	frame	K.P	P1	P2	P3	P4	P5	P6	P7
Iqra	0	903	1	**306**	202	**306**	76	12	0
Iqra	1	825	1	303	140	**306**	63	0	0
Iqra	10	889	0	304	174	**306**	84	0	0

2. range

If (B1)= then P1 weight=1
If (B2)>= cc <(B3) then P2_ weight =1
If (B3)>= cc <(B4) then P3_ weight =1

Figure 5:Result of filtering at P1 through P7

The next process is processing by eliminating duplication, normalization of weights and dominant weights so as to produce a reference table of reading conformity. Explanation of the process as follows.

4.2.4 Eleminating Duplication

The process of eliminating duplication by means of feature selection of duplicated data in each frame and readings ranging from P1 to P7 to be omitted or to 0. Figure 6 shows that there is a lot of duplication in P1 almost entirely. P2 is not entirely duplicated. Duplication process starts from P1 to P7 in each frame and reads in a way compared to one overall.

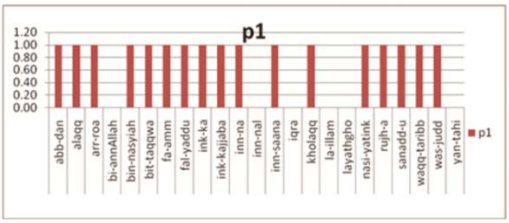


Figure 6:The overall P1 is duplicated

In Figure 6 that the P1 in frame 0 looks the same spread even if it is entirely the same starting from "abb / and" and "yan / tahi" so that the deletion is done to 0. Figure 7 shows that there is a duplication of P2 so that it is eliminated to 0. Based on the selection it eliminates the duplication of features obtained from data P2, P3, P5 and P6, whereas P1, P4 and P7 are removed because of duplication.

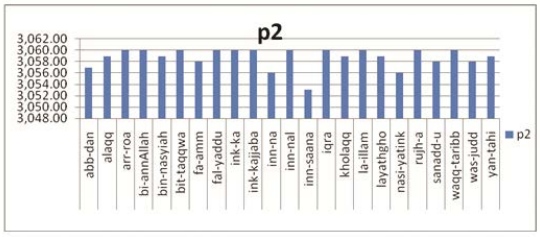


Figure 7:P2 is partly duplicated

Figure 8 shows that there is duplication so that it is eliminated and P2 becomes PF2, P3 becomes PF3, P5 becomes PF5 and P6 becomes PF6.

3.Filtering

Reading	frame	K.P	P1	P2	P3	P4	P5	P6	P7
Iqra	0	903	1	**306**	202	**306**	76	12	0
Iqra	1	825	1	303	140	**306**	63	0	0
Iqra	10	889	0	304	174	**306**	84	0	0

4. Eliminating duplication

Reading	frame	K.P	pf1	pf2	pf3	pf4	pf5	pf6	pf7
Iqra	0	903	0	0	0	0	76	12	0
Iqra	1	825	0	0	140	0	63	0	0
Iqra	10	889	0	0	174	0	84	0	0
amount		9348	0	0	1.268	0	636	80	0

Figure 8:Eliminates duplicate features from the data

Eliminating the duplication can not be a reference table of reading suitability because there is still a possibility of similarity in the process of calculation. The next process is then carried out the process of normalization.

4.2.5 Weight Normalization

The notion of normalization comes from E. F.Codd [15], one of the pioneers of database technology about relational data structure. Normalization is a technique for organizing data into tables to meet the needs of users within an organization, eliminate duplicate data by Jogiyanto [15]. The goal of normalization is to eliminate duplicate data, reduce complexity and facilitate data modification. The normalization process is normalization of weights. Normalization of weights is done by making variable amounts in advance of the variable then the number becomes a divisor. Variable sum is the sum of from each pf can be seen in Figure 9.

4. Eliminating duplication

Reading	frame	K.P	pf1	pf2	pf3	pf4	pf5	pf6	pf7
Iqra	0	903	0	0	0	0	76	12	0
Iqra	1	825	0	0	140	0	63	0	0
Iqra	10	889	0	0	174	0	84	0	0
amount		9348	0	0	1.268	0	636	80	0

Variable amount

Reading	frame	K.P	PB2	PB3	PB5	PB6
Iqra	0	903,00	0	1610361	405769	6561
Iqra	1	825,00	0	1610361	405769	6561
Iqra	10	889,00	0	1610361	405769	6561

$PB2^2 = \sum_{f0}^{f6} pf2$

Figure 9:Variable amount

Variable amount (PB) is done after completion through the process of eliminating duplication. The equation is as follows.

$$PB2^2 = \sum_{f0}^{fp} pf2 \tag{13}$$

The equation is also applicable to*PB3*^{2}*, PB5*^{2}*, PB6*^{2}

Whreas, *1) PB2 : variable amount pf2,*

2) PB3 : variable amount pf3,

3) PB5 : variable amount pf5,

4) PB6 : variable amount pf6,

5) f0: frame 0, fp: frame p.

If the variable amount (PB) has been obtained, then the process continues to normalization. The normalization used is the normalization of weights because the features of the filtered P2, P3, P5 and P6 data meet a given range given the weights so that it becomes pf2, pf3, pf5 and pf6. Next, we calculate the normalization of weights by calculating each pf2 in the frame divided by the number of PB2. The results of pf2, pf3, pf5 and pf6 have normalized weights to npf2, npf3, npf5 and npf6. The normalization equation weights as follows.

$$npf2 = \sum_{f0}^{fk} pf2 / PB2 \tag{14}$$

The equation is also applicable to *npf3, npf5, npf6*

Whereas, *1) npf2: number of normalized weights pf2, pf2: duplicated filter results P2,*

2) npf3: number of normalized weights pf3, pf3: duplicated filter results P3,

3) npf5: number of normalized weights pf5, pf5: duplicated filter results P5,

4) npf6: number of normalized weights pf6, pf6: duplicated filter results P6,

5) f0: frame 0, fk: frame k.

The following normalized weights can be seen in Table 2 on each number of features of the pre-pf comparison data and features of the npf data after normalization of weights.

Table 2:Result of normalization of weights

5.Normalization of weights										
Reading	frame	K.P	Before Normalization				After Normalization			
			pf2	pf3	pf5	pf6	npf2	npf3	npf5	npf6
Iqra	0	903	0	0	76	12	0	0	0.11	0.14
Iqra	1	825	0	140	63	0	0	0.11	0.09	0
Iqra	2	895	0	190	0	0	0	0.14	0	0
Iqra	3	848	0	158	68	12	0	0.12	0.10	0.14
Iqra	4	821	0	140	59	12	0	0.11	0.09	0.14
Iqra	5	826	0	149	58	9	0	0.11	0.09	0.11
Iqra	6	811	0	145	50	6	0	0.11	0.07	0.07
Iqra	7	858	0	172	64	11	0	0.13	0.10	0.13
Iqra	8	802	0	0	50	10	0	0	0.07	0.12
Iqra	9	870	0	0	64	8	0	0	0.10	0.09
Iqra	10	889	0	174	84	0	0	0.13	0.13	0
amount		9348	0	1269	637	81	0	0.99	1.00	0.99

Normalization of weights also can not be a reference table of reading suitability so it needs to be done next process which is the selection of dominant weight. Selection of dominant weights aims to

determine which areas are the most dominant or superior to be used as reference table reading suitability. Explanation of the dominant weight process as follows.

4.2.6 Dominant Weight

The dominant weight is the weight that becomes dominant or superior or prominent. The selection of features of the data in the form of dominant weight is taken by taking the highest in each frame in npf2, npf3, npf5 and npf6 being npf2d, npf3d, npf5d and npf6d as the dominant weights. The dominant taking is taken on each frame and readings. Table 3 comparison before and after taken dominant weight.

Table 3:Results of dominant weight

Surat Al-Alaq													
		Before dominant weight				After dominant weight							
Frame	Reading	npf2	npf3	npf5	npf6	2-dn	npf2d	3-dn	npf3d	5-dn	npf5d	6-dn	npf6d
0	Iqra	0	0	0.11	0.14	5	0	10	0	2	0.11	**1**	**0.14**
1	Iqra	0	0.11	0.09	0	4	0	7	0.11	6	0.09	11	0
2	Iqra	0	0.14	0	0	2	0	1	0.14	11	0	9	0
3	Iqra	0	0.12	0.10	0.14	11	0	4	0.12	3	0.10	**3**	**0.14**
4	Iqra	0	0.11	0.09	0.14	10	0	8	0.11	7	0.09	**2**	**0.14**
5	Iqra	0	0.11	0.09	0.11	9	0	5	0.11	8	0.09	6	0.11
6	Iqra	0	0.11	0.07	0.07	8	0	6	0.11	10	0.07	8	0.07
7	Iqra	0	0.13	0.10	0.13	7	0	3	0.13	5	0.10	4	0.13
8	Iqra	0	0	0.07	0.12	6	0	11	0	9	0.07	5	0.12
9	Iqra	0	0	0.10	0.09	1	0	9	0	4	0.10	7	0.09
10	Iqra	0	0.13	0.13	0	3	0	2	0.13	1	0.13	10	0

One example of feature retrieval from the dominant weight data in the npf6d column is 0.14. The columns contain frame 0, frame 3 and frame 4 and are ranked dominantly with the names of 6-dn ranks 1, 2 and 3. The difference lies in before using the dominant weights on npf2, until npf6 is retrieved all there is no dominant rank on each frame, while after using the feature of the dominant weight data there is a dominant rank on each frame. The dominant frame is searched by testing up to how many dominant frames to be the reference tables reading suitability. The reference table of conformity can be seen in Tables 4, 5 and 6.

Table 4:Result of limit reference table

Reference tables of Limit									
Reading	Frame	Reading Laws	K.P	B1	B2	B3	B4	B5	B6
Iqra	0	Qalqalah	903	-35.13	2.67	22.33	40.48	78.30	116.11
Iqra	1	Qalqalah	825	-18.05	13.5	28.24	45.24	76.88	108.53
Iqra	10	Qalqalah	889	-25.12	6.16	24.97	37.46	68.75	100.04
		CUP	0.4/0.5						

Table 5:Result of variable amount reference table

Reference tables of *variable* amount						
Reading	frame	K.P	PB2	PB3	PB5	PB6
Iqra	0	903.00	0	1610361	405769	6561
Iqra	1	825.00	0	1610361	405769	6561
Iqra	10	889.00	0	1610361	405769	6561

Table 6:Results of dominant weight table reference

Reference tables of dominant weight									
Frame	Reading	2-dn	npf2d	3-dn	npf3d	5-dn	npf5d	6-dn	npf6d
0	Iqra	5	0	10	0	2	0.11	1	0.14
1	Iqra	4	0	7	0.11	6	0.09	11	0
10	Iqra	3	0	2	0.13	1	0.13	10	0

The reference table consists of limits B1 to B6, CUP, overall P (K.P), variable amount PB2 through PB6, and Normalization of Dominant Weight (NDW) on each frame and reading.

After yielding reference table readiness appropriateness obtained then processed to next step that is test to input reader test and also input reading source.

5 Testing Stage

In the process of testing the conformity stages of reading the Qur'an compare the existing features in the reference table with the features available on the testers reading. Testing is not only on the testers' reading but also on the source reading. Testing stages starting from the recording of test readers then performed the MFCC process to get the features cepstral coefficient. The cepstral coefficient feature is processed by first filtering into the reference table in the form of a 1st to 6th feature limit. After going through the filtering process with the range coming from the boundary feature, then the next process performs sequential multiplication count. The calculation of the reference table in the form of feature variable number and feature of dominant weight. After the sequential multiplication counting process is obtained, then the final count calculation is the Conformity Uniformity Pattern (CUP) formula. Tests for the readiness of the Qur'an recitation are sampled against 11 Qur'anic letters, 8 reading codes and 886 recording files. Stages of testing steps are carried out as follows.

5.1 Input Testers' Readings

The reading test is done by taking the test reader input of 5-7 people on 11 letters containing 8 reading laws.

5.2 Records of Testers

Taking sound recordings of testers as many as 886 recording files. The recording of the reading is extracted with MFCC which produces the cepstral coefficient feature, which is matched to the corresponding reading to the reference table.

5.3 MFCC Readers Test

The subsequent reading sound of the testers using MFCC produces the frame and cepstral coefficient features which are then performed to the conformity testing stage.

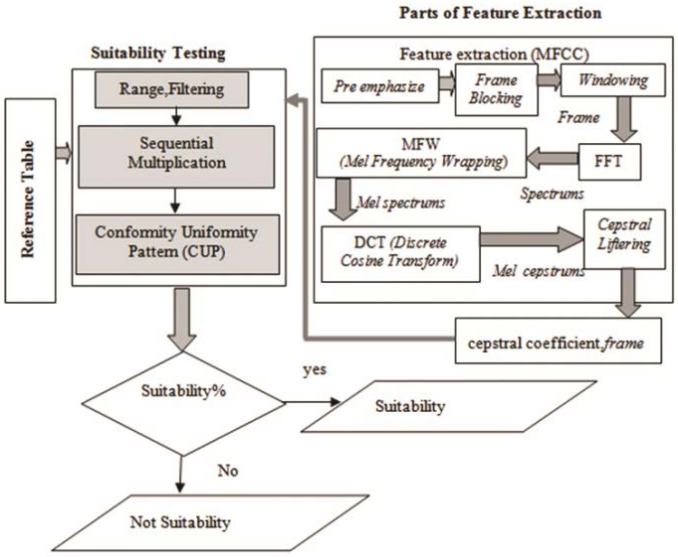


Figure 10:Section of extraction and conformity testing

The reading suitability test as shown in Figure 10 begins with range, filtering, sequential multiplication and the Conformity Uniformity Pattern (CUP). The testing process is carried out on the features in the reference reference reading table. Explanation of process range and filtering as follows.

5.4 Range and Filtering Testers

Filtering is the process of filtering testers with a certain range. Screening of test data readings starting from MFCC which produces cepstral coefficient feature then tested against the existing features in the reference table in the form of a boundary that forms the range with if it meets the conditions then given the weight equal to one. The equation is as follows.

1. $Condition\ of\ R1 is\ if\ (limit1) minimum\ value = cc\ then\ P1WeightTest = 1,$
2. $Condition\ of\ R2\ is\ if\ (((limit1) \geq= cc)\ and\ (cc < limit2)) then\ P2WeightTest = 1,$
3. $Condition\ of\ R3\ is\ if\ (((limit2) \geq= cc)\ and\ (cc < limit3)) then\ P3WeightTest = 1,$
4. $Condition\ of\ R4\ is\ if\ (((limit3) \geq= cc)\ and\ (cc < limit4)) then\ P4WeightTest = 1,$
5. $Condition\ of\ R5\ is\ if\ (((limit4) \geq= cc)\ and\ (cc < limit5)) then\ P5WeightTest = 1,$
6. $Condition\ of\ R6\ is\ if\ (((limit5) \geq= cc)\ and\ (cc < limit6)) then\ P6WeightTest = 1,$
7. $Condition\ of\ R7\ is\ if\ (limit6) maximum\ value = cc\ then\ P7WeightTest = 1\ .$ (15)

In this case, *cc: features of cepstral coefficients.*

Checking filtering is done based on existing features in the reference table is the limit and range features.

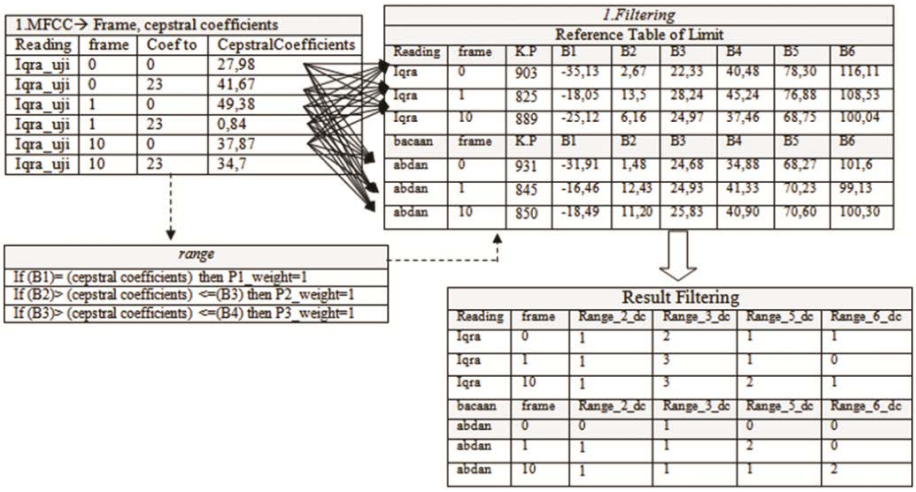


Figure 11:Range and filtering

The reference table used is the limit from B1 to B6. Figure 11 shows that the features of the frame and cepstral coefficient are skewed with the filtering features present in the limit reference table. The matching processes that are based on a range that, when fulfilled, are weighted and accumulated are stored in the range_2_dc, range_3_dc, range_5_dc and range_6_dc. The sequential multiplication process is described as follows.

5.5 Sequential Multiplication

The sequential multiplication is the process of multiplying the reading of the testers by the existing reading of the features in the reference table of the suitability of the sequential reading. The equation is as follows.

$$P2uji = \sum_{f0}^{fr} (range\_2\_dc \times npf2d \times PB2^2) \quad (16)$$

This equation is also applicable to*P3uji, P5uji, P6uji*

Whereas, *1) P2Uji: the number of P2 test,,*

2) P3Uji: the number of P3 test,

3) P5Uji: the number of P5 test,

4) P6Uji: the number of P6 test,

5) range_2_dc: filter range P2,

6) range_3_dc: filter range P3,

7) range_5_dc: filter range P5,

8) range_6_dc: filter range P6,

9) npf2d: normalization of dominant weight P2,

10) npf3d: normalization of dominant weight P3,

11) npf5d: normalization of dominant weight P5,

12) npf6d: normalization of dominant weight P6,

13) PB2: variable amount of P2,

14) PB3: variable amount of P3,

15) PB5: variable amount of P5,

16) PB6: variable amount of P6.

On the equation above, it is done one by one process of filtering result sequentially on the feature of variable amount and also to feature of reference table that is feature of dominant weight as shown in Figure 12.

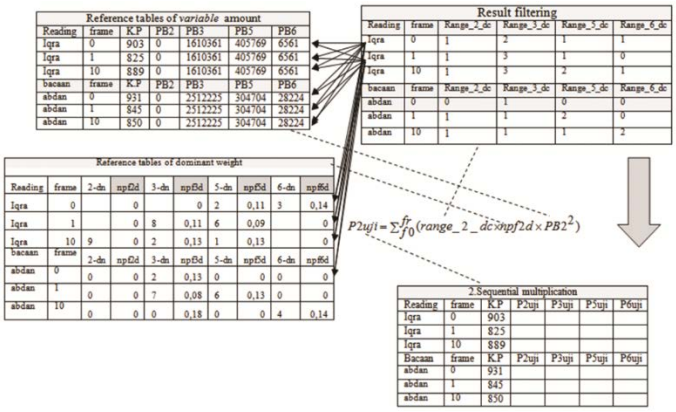


Figure 12:The sequential multiplication process of the test's reader

The sequential multiplication calculation is then processed into the Conformity Uniformity Pattern (CUP). The calculation phase of the CUP is described as follows.

5.6 Conformity Uniformity Pattern(CUP)

Conformity Uniformity Pattern (CUP) is the multiplication of the average number of P2Uji, P3Uji, P5Uji, P6Uji with the average P total (K.P) and divides by the average variable of the number raised. The reading suitability count with Conformity Uniformity Pattern(CUP) as follows.

$$CUPP2 = \frac{\overline{(P2uji)} \times \overline{(K.P)}}{\overline{(PB2^2)}} \tag{17}$$

The equation is also applicable to *CUPP2 , CUPP3, CUPP5, CUPP6*

$$CUPP2P3P5P6 = \frac{\overline{(P2uji)} + \overline{(P3uji)} + \overline{(P5uji)} + \overline{(P6uji)} \times \overline{(K.P)}}{\overline{(PB2^2)} + \overline{(PB3^2)} + \overline{(PB5^2)} + \overline{(PB6^2)}}$$

Whereas,

1) CUPP2: Conformity Uniformity Pattern P2,

2) CUPP3: Conformity Uniformity Pattern P3,

3) CUPP5: Conformity Uniformity Pattern P5,

4) CUPP6: Conformity Uniformity Pattern P6,

5) P2Uji: P2 test,

6) P3Uji: P3 test,

7) P5Uji: P5 test,

8) P6Uji: P6 test,

9) PB2: variable amount of P2,

10) PB3: variable amount of P3,

11) PB5: variable amount of P5,

12) PB6: variable amount of P6,

13) K.P: overall of P (P2, P3, P5, P6).

The CUP calculations on the features contained in the reference table are performed to measure how well the resource reads the testers' reading. The test is also conducted by measuring the level of suitability level of reading that the CUP in can from the existing features in the reference table approaching between 0.4 to 0.5 as in Table 7.

Table 7:Results of Al-Alaq's letter of CUP

frame	Reading	K.P	P1	P2	P3	P4	P5	P6
0	abbdan	931.00	1.00	306.00	210.00	306.00	83.00	24.00
1	abbdan	845.00	0.00	304.00	139.00	306.00	72.00	24.00
2	abbdan	847.00	0.00	303.00	137.00	306.00	73.00	28.00
3	abbdan	877.00	0.00	304.00	169.00	306.00	77.00	21.00
4	abbdan	887.00	0.00	304.00	169.00	306.00	80.00	28.00
5	abbdan	873.00	0.00	304.00	154.00	306.00	78.00	31.00
6	abbdan	878.00	0.00	304.00	157.00	306.00	80.00	31.00
7	abbdan	832.00	0.00	303.00	134.00	306.00	68.00	21.00
8	abbdan	906.00	0.00	304.00	179.00	306.00	83.00	34.00
9	abbdan	1,041.00	0.00	306.00	290.00	306.00	103.00	36.00
10	abbdan	850.00	0.00	304.00	144.00	306.00	72.00	24.00
	amount	9,767.00	1.00	3,346.00	1,882.00	3,366.00	869.00	302.00
Average		887.91		304.18	171.09	306.00	79.00	27.45
	square			11,195,716.00	3,541,924.00		755,161.00	91,204.00
	CUP	**0.43/0.5**		0.02	0.04		0.09	0.27

CUP between 0.4 up to 0.5 is used as reference table can be done base to do level of reading suitability testing. Tests of reading confidence with CUP can be seen in Figure 13 after the sequential multiplication in P2uji, P3uji, P5uji and P6uji are included into the Conformity Uniformity Pattern (CUP). The results of the CUP seen highest between 0.5 and 0.4 are dominant and superior, so it looks appropriate or not according to the existing reading on the reference table. CUP test with the result 0.4 looks more in accordance with the existing count on the features contained in the reference table so that the reading tested by reference is the sound of "iqra" while in reading "abbdan" the CUP result is not close to 0.4 or 0.5 so it is not according to the reading.

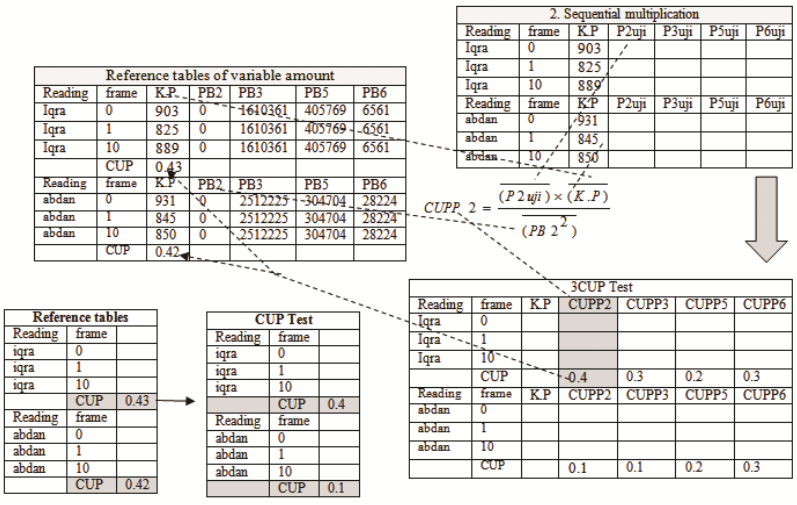


Figure 13:Counting the CUP process

CUP test taken highest 0.4 closer to the reference table of "iqra" sound, while the reading of "abbdan" at CUP Test is taken the highest 0.3 is also not included between 0.4 and 0.5

6 Test Result

The research was conducted on 11 letters, 8 reading law, and 886 recording files. The first test conducted is the test of the dominant frame consisting of 11 frames. The second test is performed on the number of cepstral coefficient, and the third test is carried out on the number of frames.

6.1 Result of Frame Dominant Testing

This test is done to find the most dominant frame area between frame 0 to 10th frame. The dominant frame test was obtained with an accuracy rate of 91.37% on the dominant 9 frames, while the dominant frame 5 (5-dm) up to 1 (1-dm) frame decreased as shown in Table 8.

Table 8:Results of Quranic reading test

No	Letter	1-dm	2-dm	3-dm	4-dm	5-dm	6-dm	7-dm	8-dm	**9-dm**	10-dm	11-dm
1	Al-Fatehah	21.43	37.01	67.53	77.92	87.01	89.61	90.91	90.91	**90.91**	90.91	90.26
2	Al-Baqarah	43.48	72.46	75.36	82.61	82.61	86.96	86.96	86.96	**86.96**	86.96	85.51
3	Al-Imran	41.67	65.28	81.94	81.94	83.33	84.72	84.72	84.72	**87.50**	87.50	87.50
4	Ar-rohman	53.85	71.79	73.08	82.05	84.62	83.33	82.05	83.33	**84.62**	84.62	84.62
5	Al-Hadid	51.52	65.15	68,18	78.79	84.85	84.85	78.79	86.36	**86.36**	86.36	86.36
6	Al-Alaq	42.75	62.32	69.57	84.78	93.48	95.65	95.65	96.38	**95.65**	96.38	96.38
7	Al-Ashr	56.36	65.45	69.09	70.91	74.55	72.73	92.73	94.55	**100.00**	100.00	100.00
8	Al-Kautsar	51.85	83.33	81.48	83.33	83.33	83.33	100.00	98.15	**100.00**	98.15	100.00
9	Al-Ikhlas	58.33	53.33	66.67	83.33	81.67	83.33	86.67	91.67	**93.33**	91.67	91.67
10	Al-Falaq	58.67	74.67	80.00	81.33	90.67	92.00	93.33	93.33	**92.00**	92.00	93.33
11	An-nas	63.08	73.85	84.62	84.62	83.08	87.69	84.62	87.69	**87.69**	86.15	86.15
	Average	49.36	65.88	74.32	81.06	84.47	85.84	88.77	90.37	**91.37**	90.97	91.07

Description dm = dominant

Testing the accuracy value against the dominant frame of 11 frames indicates that there are 9 dominant frames (9-dm) with good accuracy value this is seen as in Table 9.

Table 9:Results of 9-dm dominant reading tests

NO	Letter Al-Qur'an	Number Letter	L	P	Juz	Person	Record	**Average (9-dm)**	amount Record
1	Al-Fatehah	1	3	4	1	7	22	**90.91**	154
2	Al-Baqarah	2	3	0	1	3	23	**86.96**	69
3	Al-Imran	3	3	0	1	3	24	**87.50**	72
4	Ar-rohman	55	3	0	27	3	26	**84.62**	78
5	Al-Hadid	57	3	0	27	3	22	**86.36**	66
6	Al-Alaq	96	3	3	30	6	23	**95.65**	138
7	Al-Ashr	103	3	2	30	5	11	**100.00**	55
8	Al-Kautsar	108	3	3	30	6	9	**100.00**	54
9	Al-Ikhlas	112	2	3	30	5	12	**93.33**	60
10	Al-Falaq	113	2	3	30	5	15	**92.00**	75
11	An-nas	114	3	2	30	5	13	**87.69**	65
	amount					51	200	**91.37%**	886

Testing of 11 letters indicates that the dominant 5-dm down is getting smaller or its accuracy decreases.

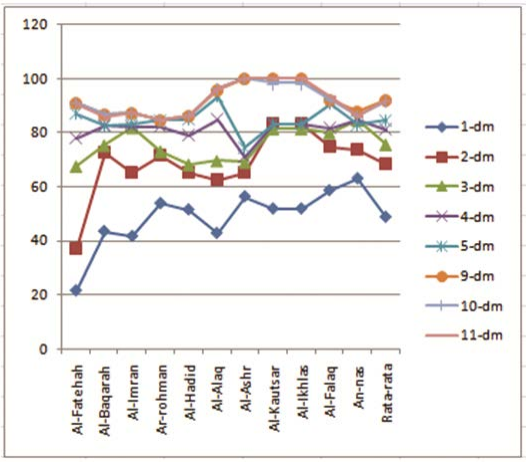


Figure 14:Comparison of the best dominant frames of 11 letters

Figure 14 shows that the dominant taking is preferably above 5-dm as it has an increased accuracy rate.

6.2 Test Results Number of Cepstral Coefficient

Tests were also performed on the amount of cepstral coefficient. Taking the number of cepstral coefficient with c-23 shows a good accuracy level with an average of 96.65%.

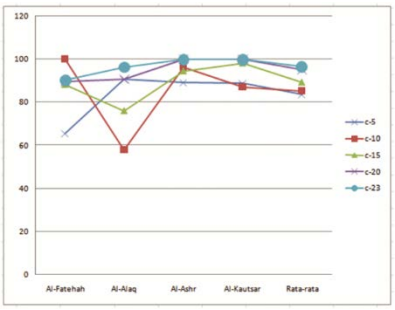


Figure 15:Comparison of the best amount of cepstral coefficient

Figure 15 shows that the samples of the five letters of Al-Fatehah, Al-Alaq, Al-Ash, Al-Kautsar show the highest on cepstral coefficient 23 or c-23 with an average of 96.65%.

6.3 Test Results Number Frame

Tests on the number of frames also performed starting from the frames F-5, F-7, F-9 and F-10. The test was conducted to find the number of frames that have the best accuracy on F-10 with an average of 96.65%. Figure 16 shows that the F-10 frame has the highest average compared to F-5, F-7 and F-9.

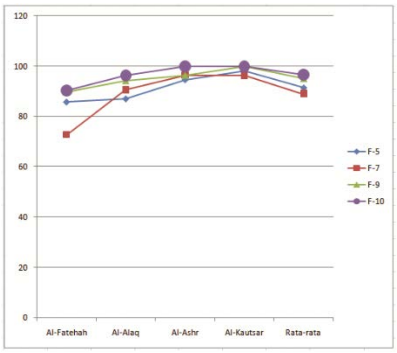


Figure 16:Comparison of the best frame counts

Based on the CUP testing, the dominant frame rate is found in the dominant 9 frames, the test for the number of cepstral coefficients in the can be c-23 with an average of 96.65%, while the best frame number in F-10 with an average of 96.65 %. Tests with the CUP formula can also determine the suitability of reading uniformity, which is contained in the reference table with CUP 0.4 - 0.5. Taking the source reading before the reference can be known feasibility if less than 90% of the source reading test should be re-recorded and re-reference.

7 Conclusion

Correct reading of the Qur'an is a problem that requires its own challenge, tailored to the rules. Research on reading conformity is done with the features in the reference table of reading fit, test and readability accuracy. Tests with the CUP calculations performed on dominant frames resulted in an average accuracy rate of 91.37% at 9 frames (9-dm). The test was also conducted on the number of cepstral coefficient to produce the best cepstral coefficient c-23 with an average of 96.65%, while testing the number of frames yielded the best F-10 with an average of 96.65%. This research yields some contribution that is:

1. The first contribution that the model extraction concept extraction uses MFCC and NDW model can be used to generate the reference table of Al-Qur'an reading conformity with the accuracy level of 91.37%. The retrieval of the reference table on the dominant frames suggested above 5 (5-dm).
2. The second contribution to the test model based on the level or level of reading conformity with the Conformity Uniformity Pattern (CUP) of 0.4 with the high level to 0.5 with the low level can know the ability of reading.

The third contribution in the process of testing the source reading before the reference table can be known the feasibility of reading the source is declared in accordance uniformity to be a reference table.

REFERENCES

[1] Zarkasyi I, "Pelajaran Tajwid," *Gontor Ponorogo*, vol. Trimurti P, p. Hal 1-3, 1995.

[2] Suyanto and S. Hartati, "Design of Indonesian LVCSR using Combined Phoneme The Approaches of LVCSR," *Icts*, pp. 191–196, 2013.

[3] S. Suyanto and A. E. Putra, "Automatic Segmentation of Indonesian Speech into Syllables using Fuzzy Smoothed Energy Contour with Local Normalization, Splitting, and Assimilation," *J. ICT Res. Appl.*, vol. 8, no. 2, pp. 97–112, 2014.

[4] R. Cahyarini, U. L. Yuhana, and A. Munif, "Rancang Bangun Modul Pengenalan Suara Menggunakan Teknologi Kinect," *J. Tek. Pomits*, vol. 2, no. 1, pp. 1–5, 2013.

[5] M. L. Chen, S. K. Changchien, X. M. Zhang, and H. C. Yang, "The design of voice recognition controller via grey relational analysis," *Proc. 2011 Int. Conf. Syst. Sci. Eng. ICSSE 2011*, no. June, pp. 477–481, 2011.

[6] M. Bodruzzaman, K. Kuah, T. Jamil, C. Wang, and X. Li, "System Using Artificial Neural Network," pp. 1–3, 1993.

[7] B. P. Tomasouw and M. I. Irawan, "Multiclass Twin Bounded Support Vector Machine Untuk Pengenalan Ucapan," *Pros. Semin. Nas. Penelitian, Pendidik. dan Penerapan MIPA, Fak. MIPA, Univ. Negeri Yogyakarta,* vol. 2, no. 2004, pp. 1–10, 2012.

[8] M. A. Pathak, "Privacy-Preserving Machine Learning for Speech Processing," *PhD*, pp. 1–140, 2012.

[9] A. M. Aibinu, M. J. E. Salami, A. R. Najeeb, J. F. Azeez, and S. M. A. K. Rajin, "Evaluating the effect of voice activity detection in isolated Yoruba word recognition system," *2011 4th Int. Conf. Mechatronics Integr. Eng. Ind. Soc. Dev. ICOM'11 - Conf. Proc.*, no. May, pp. 17–19, 2011.

[10] S. Hidayat, R. Hidayat, and T. B. Adji, "Sistem Pengenal Tutur Bahasa Indonesia Berbasis Suku Kata Menggunakan MFCC, Wavelet Dan HMM," *Conf. Inf. Technol. Electr. Eng.*, no. September, pp. 246–251, 2015.

[11] S. B. Davis and P. Mermelstein, "Comparison of Parametric Representations for Monosyllabic Word Recognition in Continuously Spoken Sentences," *IEEE Trans. Acoust.*, vol. 28, no. 4, pp. 357–366, 1980.

[12] K. Chakraborty Asmita Talele Savitha Upadhya, "Voice Recognition Using MFCC Algorithm," *Int. J. Innov. Res. Adv. Eng.*, vol. 1, no. 10, pp. 2349–2163, 2014.

[13] H. S. Manunggal, "Perancangan dan Pembuatan Perangkat Lunak Pengenalan Suara Pembicara Dengan Menggunakan Analisa MFCC Feature Extraction.," *Tugas Akhir Sarj. pada Jur. Tek. Inform. Fak. Teknol. Ind. Univ. Kristen Petra Surabaya*, 2005.

[14] A. E. Putra, "Frekuensi Cuplik pada FFT," *Tan Li, Process. Digit. Signal*, vol. 1, 2008.

[15] J. E.F Codd, P. Ritonga, and L. A. Reply, "Pengertian Normalisasi Database Dan Bentuk-," pp. 3–5, 2015.

AIVP ADVANCES IN IMAGE AND VIDEO PROCESSING
SOCIETY FOR SCIENCE AND EDUCATION UNITED KINGDOM
ISSN 2054-7412

No-Reference Image Quality Assessment Based on Edges

[1]Chun-Chieh Chang, [2]Chin-Chen Chang
[1,2]Department of Computer Science and Information Engineering, National United University, Miaoli 360, Taiwan;
ccchang@nuu.edu.tw

ABSTRACT

Image quality assessment is a crucial topic in the field of image processing. In this paper, we propose an edge-based no-reference image quality assessment method. The following factor is applied to assess image quality, namely, improved blur measurement. In the improved blur measurement method, we propose an algorithm that improves the accuracy in measuring image blurs and attains effective execution speed in time complexity. Experimental results reveal that using the proposed approach helps attain satisfactory image quality assessment results.

Keywords: Image Quality; Blurring; No-reference; Image Edge.

1 Introduction

Because of the demands for and popularity of various types of imaging equipment, improving image quality analysis abilities to enable images precisely reflect the characteristics of the objects they portray has become the most crucial objective of image quality measurements. In modern image processing studies, the problem of image distortion during the capturing, transmission, and compression processes can be solved by using assessment indicators, which assess image quality and the degrees of distortion and blur. This enables higher-quality images to be obtained for use in the relevant equipment or systems.

Studies on the image quality technologies employed for computer vision have utilized objective computer digital technologies to capture and obtain image content similar to that observed by the human eye. Therefore, identifying how digital technologies can be used to capture images that are deemed authentic and clear by human observers is the ultimate goal of image quality–related studies. For example, when developing image quality technologies, image detection, identification, and analysis technologies are critical approaches to identify and address problems such as inadequate lighting, blurred images, exposure problems, and poor image quality (for photographs taken by digital cameras). The human visual system (HVS) presents an intuitive understanding of image perception. By contrast, blurring can be caused by problems with digital equipment, such as errors created by objects moving during the image capturing process, errors in the digital or analogue quantification process, changes created by object edges shifting, and edge signal blurring created by digital equipment when attempting to reduce noise. When these problems occur, digital equipment produces blurred images. Therefore, regardless of the blur type (e.g., dynamic blur or out-of-focus blur), strengthening the image information for digital equipment and designing high-quality algorithms to elevate interpretation accuracy are required for image quality measurements. [2, 3, 6]

Among the image quality analysis and measurement methods, the no-reference (NR) image quality assessment method may be used in all operating environments. However, because this method does

DOI: 10.14738/aivp.62.4385
Publication Date: 27th April 2018
URL: http://dx.doi.org/10.14738/aivp.62.4385

not utilize original images for references and comparisons, measurements are prone to large errors. Nevertheless, the goal of the NR method is to develop a system that is comparable to the HVS. Image distortion–related studies, such as those on blur, the blocking effect, and sharpness, have produced favorable results [5, 12–13]. The blur measure (BM) method [4] is an algorithm based on the NR image quality assessment method, in which the Sobel algorithm is used to detect image edges. In the Sobel measurement method, one-dimensional analysis is first used to identify the pixel with the largest change in edge color, after which two-dimensional analysis is conducted. In this paper, an improved blur measure (IBM) was proposed in order to develop a more sophisticated algorithm for measuring the degree of blur. In the proposed algorithm, improvement strategies were implemented to increase performance and strengthen image operation speed and capabilities, thus enabling blurred images to be accurately measured and used.

The rest of this paper is structured as follows. Section 2 reviews related works. Section 3 describes our method. Section 4 describes the experimental results. Finally, Section 5 presents the conclusions.

2 Related Works

Objective image quality assessments involve classifying original and processed images into categories and determining changes in image distortion through comparison. The assessment results are influenced by factors including tone reproduction, detail and edge reproduction, noise, and color reproduction. Objective image quality assessments can be broadly divided into three categories: full-reference (FR) [7], reduced-reference (RR) [14], and NR image quality assessments [1, 8]. FR image quality assessments compare the differences in the pixels of complete images before and after image processing to identify changes in the degree of distortion between original and processed images. This type of assessment is suitable for subjective assessments of image quality with the support of objective assessment indicators that are intended to yield optimal-quality images. RR image quality assessments extract and compare the eigenvalues before and after image processing to reveal the attenuation situation. This type of assessment is a combination of FR and NR image quality assessments. NR image quality assessments do not require original images, and they determine image quality by examining the error results after image processing. Thus, the calculation and comparison costs of NR image quality assessments are substantially lower than those of FR and RR image quality assessments, rendering the NR methods frequently studied.

The blur measurement method [4] mainly involves using Sobel operators to measure image edges, and the measurement results are compared with those obtained using the just noticeable blur (JNB) method. This method yields higher difference of mean opinion scores. The blur measurement method uses the Sobel method for image edge detection. The Sobel method first performs one-dimensional analysis to identify pixels with the maximum edge color changes, and proceeds to two-dimensional analysis. The Sobel method employs a connected-component method that adopts the concept of image pixel neighbors to connect adjacent pixels.

Calculation models defined using the Sobel method yield more accurate results than those obtained using the JNB method. Thus, compared with the JNB method, the aforementioned computer vision calculation models produce blurred image identification results closer to those obtained using the HVS. For results obtained using the blur measurement method, a higher Gaussian blur standard deviation indicates a lower blur value, signifying a higher degree of blur. By contrast, a lower Gaussian blur standard deviation indicates a higher blur value, signifying a smaller degree of blur and a clearer image.

3 The Proposed Approach

This paper introduced an improved blur measurement calculation model called the IBM. Because humans are more sensitive to edge intensity changes when images are clear, changes in pixel intensity between neighboring areas are higher in high-quality images (e.g., clear images) than in low-quality images (e.g., blurred images). The IBM was developed on the basis of this principle. In IBM calculations, a lower value indicates a higher degree of blur and poorer object visibility in the image, whereas a higher value indicates a lower degree of blur and more favorable visibility.

The IBM was developed on the basis of pixel intensity changes around image edges. High-quality clear images contain less blur and high pixel intensity changes around image edges, whereas blurred and distorted images contain low pixel intensity changes around the edges. Horizontal and vertical detections were performed to calculate the number of edge pixels in the original. In addition, the formula was further modified to generate normalized numbers to observe changes in blurred images more clearly. The revised formula is shown as follows:

$$IBM = \frac{\sum_{I(x,y)\in HE}\sqrt{\sum_{I(x',y')\in VN_{xy}}\{I(x,y)-I(x',y')\}^2/\left|VN_{xy}\right|}}{\sqrt{\sum_{I(x,y)\in HE}I(x,y)}} + \frac{\sum_{I(x,y)\in VE}\sqrt{\sum_{I(x',y')\in HN_{xy}}\{I(x,y)-I(x',y')\}^2/\left|HN_{xy}\right|}}{\sqrt{\sum_{I(x,y)\in VE}I(x,y)}} \quad (1)$$

where HE and VE are defined as the horizontal and vertical edge pixels, respectively. HNxy and VNxy are defined as the neighboring horizontal and vertical pixels of pixel I(x,y), respectively, where I(x,y) ϵ HE and VE, and |HNxy| and |VNxy| are the numbers of pixels of HNxy and VNxy, respectively.

In (1), a higher value denotes higher image definition and more changes in edge pixel intensity, whereas a lower value denotes a higher degree of blur and fewer changes in edge pixel intensity. The formula introduced was developed according to the concept of blur measurement in order to clarify the differences between clear and blurred images. The original blur measurement method was revised to generate larger values that represented greater differences. Thus, high-quality clear images contained less blur and more pixel intensity changes around image edges. By contrast, blurred images contained fewer pixel intensity changes around image edges.

Distorted images were created through blurring, after which NR image quality assessments were performed; the differences between the original images and the distorted images were subsequently assessed. To create distorted images, the Gaussian blur processing method was adopted. Gaussian blur, also known as Gaussian smoothing, is a type of low-pass filter and a function that is widely used in image processing software. Gaussian blur is commonly used to remove noise and diminish detail level to facilitate image recognition. In this paper, original images were processed using the Gaussian blur method with various standard deviations to create blurred images, which then formed the image database used in the experiment. Gaussian blur uses normal distribution to obtain pixel values and calculate the value changes of all pixels in an image. Therefore, the smaller the distance between a pixel and the blur center, the higher the weight value becomes. Compared with the mean filter method, the Gaussian blur method features more effective smoothing and better edge effect retention.

4 Results

In the experiments, Windows 10 64-bit was used as the operating system and programs were written using MATLAB 2014. An Intel(R) Core(TM) i7 CPU 930 @ 2.80GHz with a memory of 12 GB was used as the CPU. The TID2013 image quality assessment database [9–11] was downloaded from the Internet to test the proposed method. The 125 blurred images in the TID2013 database were created by blurring 25 high-definition RGB images (Figure. 1) using Gaussian blurring with various standard deviations. A mean opinion score was assigned to each image in the database. Fig. 2 shows the images created using varying degrees of Gaussian blurring, where Images 1 and 5 were the clearest and most blurred images, respectively.

Figure. 1. Images used in the experiment

Original image

Image 1

Image 2

Image 3

Image 4 **Image 5**

Figure. 2. Degrees of Gaussian blurring shown in image database images

4.1 Image Quality Measurements

Three images from the image database were used to test the IBM, as shown in Figs. 3–5. The y-axis represents values obtained using the IBM, and the x-axis represents the image number; a larger image number represents a greater degree of Gaussian blurring. The results revealed that values obtained using the IBM were correlated with image quality. Higher and lower IBM values signified clearer and more blurred images, respectively.

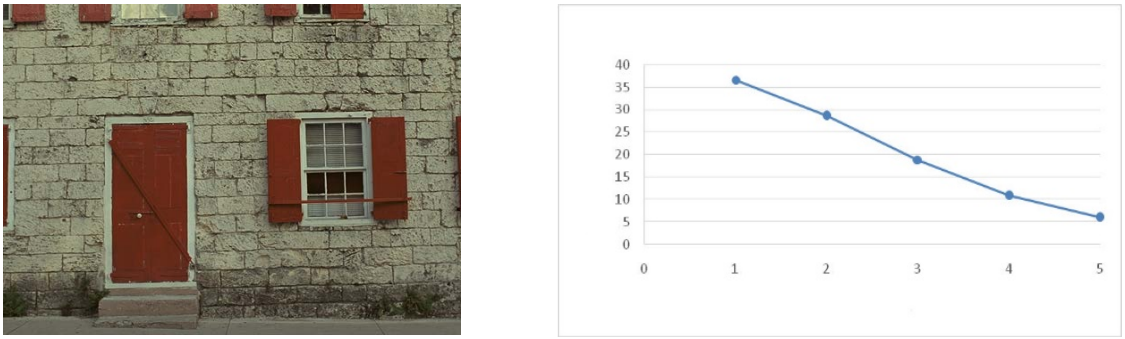


Figure. 3. IBM measurement results

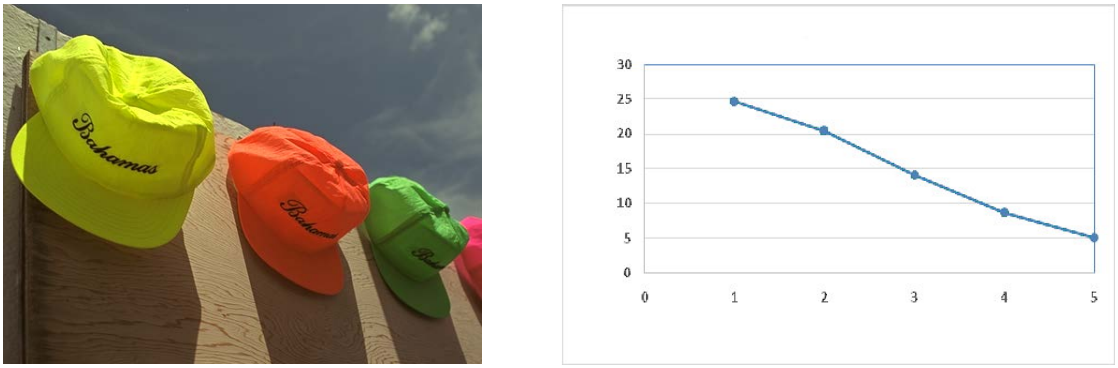


Figure 4. IBM measurement results

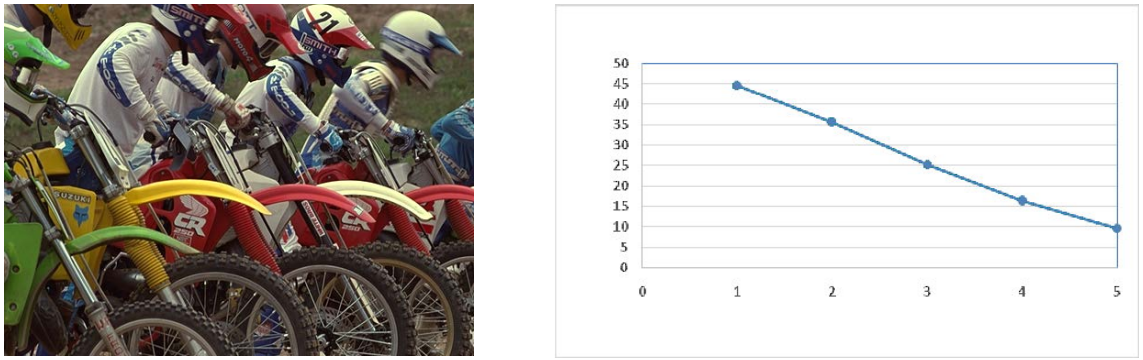


Figure. 5. IBM measurement results

4.2 Analyses of the Subjective and Objective Assessments

Correlation coefficients (CC) are commonly applied for statistical analysis. The theorem involves calculating two covariances and subtracting the standard deviations of their distributions from the two covariances. Root mean squared error (RMSE) and mean absolute error (MAE) are the most basic and widely used methods for representing random errors. However, neither RMSE nor MAE measure error values or ranges; they perform estimations on the reliability of a set of measured data.

The differences between clear and blurred images were evaluated using the proposed image quality measurement method. Subjective and objective scores were obtained and analyzed to determine whether the image quality measured was consistent with the subjective scores. Mean scores provided by the TID2013 image quality assessment database were used as the subjective scores. To verify the feasibility of the image quality assessments, three commonly used performance indicators were employed to determine the correlation between the subjective and objective scores. The CC was utilized to determine whether the image quality measurement results were consistent with the subjective scores, whereas the RMSE and MAE were used to evaluate errors and accuracies of the objective scores. Fig. 6 shows a scatter plot of the IBM results for the TID2013 database and the subjective scores, where the x-axis is the subjective score and the y-axis is the objective assessment score.

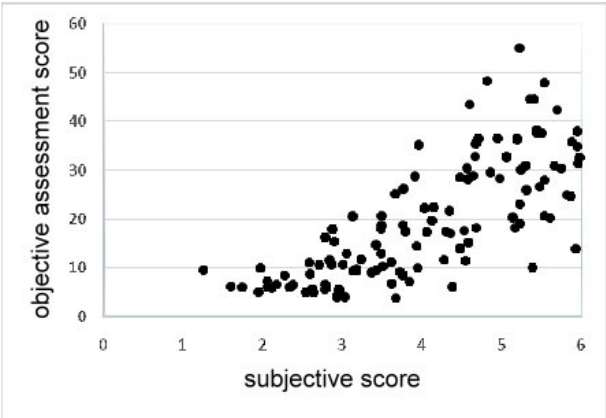


Figure. 6. Scatter plot of IBM measurements and subjective scores

The proposed image quality measurement method, subjective score scatter plot, and subjective–objective score correlation were compiled on the basis of the experimental results (Table 1). According to the image quality measurement method, the subjective and objective scores were moderately correlated.

Table 1. Correlations between objective and subjective scores

	CC	RMSE	MAE
IBM	0.7264	16.671	15.863

5 Conclusion

In this paper, an edge-based NR image quality assessment method was introduced to measure the differences between clear and blurred images. An improved algorithm was proposed to measure the degree of blur in images more accurately, promptly, and efficiently. This method also yielded favorable image quality assessment results.

REFERENCES

[1] J. Caviedes and S. Gurbuz, "No-reference sharpness metric based on local edge kurtosis," Proceedings of 2002 International Conference on Image Processing, vol. 3, pp. 53-56, 2002.

[2] D.M. Chandler, "Seven challenges in image quality assessment: past, present, and future research," ISRN Signal Processing, 2013.

[3] C.C. Chang and C.C. Chang, "An improved method for no-reference image quality assessment," International Conference on Information Technology and Industrial Application, April 2016.

[4] K. De and V.Masilamani, "A new no-reference image quality measure for blurred Images in spatial domain," Journal of Image and Graphics, vol. 1, no.1, pp. 39-42, 2013.

[5] J. Dijk, M.van Ginkel, R.J. van Asselt, L.J. van Vliet, and P.W. Verbeek, "A new sharpness measure based on Gaussian lines and edges," Proceedings of International Conference on Computer Analysis of Images and Patterns (CAIP), LNCS, vol. 2756, pp. 149-156, 2003.

[6] F.S. Frey and J.M. Reilly, Digital Imaging for Photographic Collections: Foundations for Technical Standards (Rochester, NY: Image Permanence Institute, Rochester Institute of Technology), pp.10, 1999.

[7] ITU-R Recommendation BT.500-10. Methodology for the subjective assessment of the quality of the television pictures, 2000.

[8] P. Marziliano, F. Dufaux, S. Winkler, and T. Ebrahimi, "A no-reference perceptual blur metric," Proceedings of 2002 International Conference on Image Processing, pp. 57-60, NY, USA, 2002.

[9] N. Ponomarenko, L. Jin, O. Ieremeiev, V. Lukin, K. Egiazarian, J. Astola, B. Vozel, K. Chehdi, M. Carli, F. Battisti, C.-C. Jay Kuo, Image database TID2013: Peculiarities, results and perspectives, Signal Processing: Image Communication, vol. 30, pp. 57-77, Jan. 2015.

[10] N. Ponomarenko, O. Ieremeiev, V. Lukin, K. Egiazarian, L. Jin, J. Astola, B. Vozel, K. Chehdi, M. Carli, F. Battisti, C.-C. Jay Kuo, Color Image Database TID2013: Peculiarities and Preliminary Results, Proceedings of 4th Europian Workshop on Visual Information Processing EUVIP2013, Paris, France, June 10-12, pp. 106-111, 2013.

[11] N. Ponomarenko, O. Ieremeiev, V. Lukin, K. Egiazarian, L. Jin, J. Astola, B. Vozel, K. Chehdi, M. Carli, F. Battisti, C.-C. Jay Kuo, A New Color Image Database TID2013: Innovations and Results, Proceedings of ACIVS, Poznan, Poland, pp. 402-413, October 2013.

[12] J. Tang, E.Peli, and S. Acton, "Image enhancement using a contrast measure in the compressed domain," IEEE Signal Processing Letters, vol. 10, no.10, pp. 289-292, 2003.

[13] H. Tong, M. Li, H. Zhang, C. Zhang, J. He, and W.Y. Ma, "Learning no-reference quality metric by examples," Proceedings of the 11th International Multimedia Modelling Conference, pp. 247-254, January 2005.

[14] Z. Wang and E.P. Simoncelli, "Reduced-reference image quality assessment using a wavelet-domain natural image statistic model," Proceedings of SPIE-IS and T Electronic Imaging - Human Vision and Electronic Imaging X, vol. 5666, San Jose, CA, January 2005.

AIVP ADVANCES IN IMAGE AND VIDEO PROCESSING
SOCIETY FOR SCIENCE AND EDUCATION UNITED KINGDOM
ISSN 2054-7412

Computer-Aided Rehabilitation for the Carpal Tunnel Syndrome using Exergames

[1] Ioannis Pachoulakis, [2]Diana Tsilidi, [3]Anastasia Analyti
[1,2]Department of Informatics Engineering, Technological Educational Institute of Crete Heraklion, Crete, Greece;
[3]Institute of Computer Science, Foundation for Research and Technology – Hellas (FORTH) Vassilika Vouton, Heraklion, Crete, Greece;
ip@ie.teicrete.gr; dtsilidi@gmail.com; analyti@ics.forth.gr

ABSTRACT

Carpal tunnel syndrome (CTS) has reached epidemic proportions as the surgical release of the transverse carpal ligament is included in the top ten most common operations, which significantly affects health care costs. Additional costs to consider include time off work, lost wages, and diminished workplace productivity. Non-surgical management of CTS such as splinting, non-steroidal anti-inflammatory medication, steroid injections and ergonomic modification of the work habits can help in early cases, but can be ineffective in more advanced cases, often leading to recommendation for surgical treatment. It is possible to attain meaningful physiotherapy results through interesting and engaging computer-based games which provide the motivation to continue therapy at home, even away from direct therapist supervision. Motivation and immersion in the game scenario also helps patients forget that they are performing exercises as a part of their therapy. Accordingly, we present a Unity3D Fly-A-Plane game, the scenario of which combines CTS-specific physiotherapy exercises in a natural game scenario to fly an airplane through a sequence of hoops in the sky. The game employs the Leap Motion sensor, whose detailed wrist and hand (including fingers) tracking abilities make it an excellent hardware platform for rehabilitation oriented exercises intended for patients suffering from CTS.

Keywords: Carpal tunnel syndrome; Traditional Physiotherapy; Computer-aided rehabilitation; Game-based Rehabilitation.

1 Introduction

Carpal tunnel syndrome (CTS) is a very common and frequent focal peripheral neuropathy in which the median nerve is compressed where it passes through the wrist. CTS commonly causes swelling, pain, tingling and loss of strength in the wrist and hand. Usually the tendons of the wrist become swollen and compress the nerve. The functionality of the median nerve is to control some of the muscles that make the thumb move and transfer sensory information from the thumb and fingers back to the brain [1]. Compressing the median nerve may feel painful, aching, tingling or it may cause numbness in the affected hand. During the night hours the symptoms may deteriorate the quality of the sleep and can become more obtrusive in the morning hours, after waking up. Aching and tingling can be temporarily relieved by hanging the hand out of bed or shaking it around.

Activities like typing on the computer, handwriting or some types of housework may intensify the symptoms, even if the problem is not so perceptible during the day. In some cases CTS symptoms may

DOI: 10.14738/aivp.62.4458
Publication Date: 29th April 2018
URL: http://dx.doi.org/10.14738/aivp.62.4458

resemble those for other conditions, such as arthritis or disc problems [2], a confusion that can be alleviated with the help of a nerve conduction tests. The syndrome is three times more common in women compared to men, while it affects almost 3% of the adult population. Persons who tend to perform repetitive wrist movements (especially true in the developed countries) are more prone to develop CTS. For example, CTS-related costs in the U.S. mount to over 2 billion annually, partially due to the high cost of the medical treatment for roughly 3.7% of the general public who suffer from it but also due to the missed work days. By comparison, in France, 127,269 patients over 20 years were operated to treat CTS in 2008. The syndrome is reported to affect more frequently mostly females between 45 and 59 years old, and also mostly females between 75 and 84 years old [1]. Many people, who use a computer daily (for work, education or entertainment), complain about hand paresthesia, a superficial sensation of tingling, burning or numbness of the hand [3]. However, a small percent of these computer users actually meet the clinical criteria for CTS.

Many CTS patients opt for physiotherapy, which involves tendon gliding of the finger flexor tendons and nerve gliding of the median nerve, thereby helping to reduce pain and increase mobility and strength. Additional exercises may be included, depending on the individual case, to increase muscle strength in the hand, fingers and forearm as well as in the trunk and postural back muscles, combined with stretching exercises to improve flexibility in the wrist, hand and fingers. Physiotherapy mainly aims at moderating the CTS-related symptoms in order to allow the patient to maintain the functionality of his hands in everyday activities and obviate the need for surgery. Even if surgery cannot be prevented, post-surgery physical therapy helps to restore strength and full functionality to the hand.

However, the benefits of physical therapy are not instant, but are rather felt in the long-term (many weeks or months). Daily repetition of the prescribed exercises is required to have partial recovery of impaired movement. These sessions tend to be long, tiresome and tedious and may be difficult to schedule in frequent sessions because of healthcare costs [4]. Outpatient therapy gives a solution to this problem, as the patients may perform exercises at home on a regular basis, while communicating remotely with the therapist, so that a physical therapist may not be required to be present after the first few sessions.

Even at home, treatment requires time-consuming and repetitive exercises that are not very pleasant for the patients and end up monotonous and boring, resulting in only 31% of people performing them as recommended [5]. In addition, since the presence of the physical therapist is not necessary for the performance of the exercises, it is natural for the patients to lose their motivation and confidence, while doing them alone at home, especially without immediate significant improvement to be expected. These factors allow many patients to neglect exercising on a regular basis and abandon the treatment before it is completed. Therefore, it is important to provide patients who suffer from motor disabilities rooted in neurological and musculoskeletal conditions appropriate and engaging rehabilitation programs designed to improve their quality of life. A number of studies [5], [6], [7], [8], conclude that interactive games have the potential, through several of their elements, to enhance the results of rehabilitation and therapy.

The idea of attaining meaningful physiotherapy through interesting and engaging games is promising, since it can provide the motivation to continue therapy at home, even when they are not under direct therapist supervision. Accordingly, serious games have started to be used in physiotherapy and rehabilitation, as they can promote healthy habits and exercise, while entertaining and rewarding the players [9]. Motivation and immersion in the game scenario also helps patients forget that they are performing exercises as a part of their therapy. The key factor is that exergames can employ immersive game scenarios to promote repetitive actions. In addition, goal-oriented and highly interactive games,

can provide sufficient motivation to perform and complete their exercises at home. Typical game features like controlling 3D worlds and elements like high scores, achievements, rewards and positive feedback can make therapy fun and addictive and help maintain interest and, as a result, play more and help to speed up recovery. In doing that, exergames create an environment where repetition is an advantage and something to be enjoyed rather than dreaded [10].

The present work discusses the key advantages of the Leap Motion sensor that make it a promising tool on which to base the creation of CTS-oriented serious games. These advantages have been explored through custom CTS-oriented interactive serious games which require users to perform CTS-specific exercises (extension/flexion and radial/ulnar deviation) to successfully complete the game objectives.

2 Traditional CTS Physiotherapy

When the wrist and the fingers are extended the median nerve is displaced farthest under the transverse carpal ligament into the hand. The study of McLellan and Swash [11] described how the median nerve slides distally by 10-15 mm relatively to a fixed bony landmark in the carpal tunnel, when the wrist is hyper-extended, while the displacement mentioned above, which occurs during the flexion of the wrist and the fingers, is often four times greater at the wrist that in the arm. Additionally, the nerve is getting displaced 11 mm distally during the active and passive extension of the wrist and the fingers and 4 mm proximally during the flexion. Local stretching of the median nerve, that would normally occur, is prevented by the longitudinal sliding. Szabo et al. [12] have also shown the differences between median nerve and digital flexor tendon excursion in the carpal tunnel and the linear relationship that exists between median nerve movement and that of the flexor tendons. Recurrent active wrist and digital flexion and extension exercises reduce the pressure within the carpal tunnel. Thus, it is possible to affect the clinical course of CTS of some patients by using a specific series of exercises. A systematic procedure, which actively forces the median nerve and the flexor tendons to their maximal excursion through the carpal tunnel, is proposed in case that the attachment to the median nerve exists preoperatively. Certain studies have also proposed a regimen suitable for postoperative nerve gliding that is used to minimize scar adhesions and maximize nerve excursion through the carpal canal, which is reported to be effective in the management of postoperative CTS, with relief of pain and low recurrence rates. The performance of these nerve and tendon gliding exercises significantly affects symptom resolution by stretching the adhesions in the carpal canal, broadening the longitudinal area of contact between the median nerve and the transverse carpal ligament, also by reducing tenosynovial edema by a "milking action", thus improving venous return from the nerve bundles and finally reducing pressure inside the carpal tunnel. This alternative conservative method of CTS management reduces the need for a surgical intervention.

These neuromobilization techniques may be applied when there is certain evidence of an entrapment of the median nerve at the wrist and the neurodynamic test is positive. These techniques combine repetitive movements of the hand that provoke the symptoms and movements in the more distal and proximal segments. The position and the movements necessary for the neurodynamic test are similar to those needed for the treatment exercises, which usually include movements performed in a moderate pace divided in 3 sets of 10 repetitions in each set and a 3 second hold at the final stretched position. The therapist can decide to focus on a specific part of the nerve, if the test shows that this part is responsible for the symptoms, but without abandoning the exercises for the other parts of the nerve. Especially in the first few sessions the patients may experience a feeling of ‘‘stretching’’, tissue tension, light numbness, even a slight increase of pain symptoms during the technique; however the symptoms decrease or completely stop immediately following the procedure. The patient has to

perform daily home exercise program including self-mobilization techniques for the median nerve, which can be performed in standing or sitting position.

Determining whether the condition of CTS is mild, moderate, or severe can clarify and restrict intervention choices. Before applying the specific therapy it is important to assess the severity of the symptoms, as a patient with a mild CTS condition may have a successful symptom reduction by only using splints at night and by altering the activities during the day, while a patient with moderate or severe CTS will not usually benefit so much from splint usage. Recent practice links the individual differences in CTS symptoms or the results of therapy to a genetic predisposition for developing the syndrome. As a result, people with the same activity or the same treatment might have different symptoms and treatment results.

The most common exercises recommended for the conservative management of the symptoms of CTS are tendon gliding of the finger flexor tendons and nerve gliding of the median nerve [13]. Studies that included measurements of the carpal tunnel pressure have shown that occasional exercise of active wrist and finger motion lasting about one minute lowers the pressure within the carpal tunnel [14]. The gliding exercises include moving the fingers in a specified pattern of exercises and help the tendons and nerves glide more smoothly through the carpal tunnel. While there's some evidence that gliding exercises can help relieve symptoms when used alone, these exercises appear to work better in combination with other treatments, such as splinting. Also manual therapy integrates a patented form of instrument-assisted soft tissue mobilization that gives the opportunity to the therapists to accurately detect and treat scar tissue and restrictions that effect normal function.

3 Computer-aided, CTS-oriented exergames

Serious games, a special category of games, are commonly used in training in the military, health care or engineering, but can be also a very useful tool for education. Since the games usually involve challenges and rewarding systems, they can motivate users to perform a number of tasks for the purpose of the game they would not normally undertake on their own. Accordingly, serious games find applications in health care and help mitigate severe mobility impairments due to neurological and musculoskeletal problems in multiple sclerosis and stroke sufferers. In addition, game-based physiotherapy affords patients that require regular physical therapy sessions to regularly practice at home and keep their interest, especially when results are not prompt. Accordingly, it is important to build exergames that both motivate patients [15] and facilitate the completion of repetitive, tedious tasks efficiently, so as to remain dedicated to the program and enjoy better recovery prospects [16]. In the case of CTS, the commonly prescribed gliding exercises, which are repetitive movements of the wrist and fingers can be successfully implemented in an engaging game environment with many interesting features keeping the patient amused and motivated to attain higher levels of performance.

The last decade has seen a number of attempts to take advantage of the emerging motion sensor technology in the fields of physiotherapy and rehabilitation [17], [18], and [19]. Although approaches based on academic research [20], [21] present game prototypes using sensors for physiotherapy treatment, there are also commercially available products. VirtualRehab-Hands [22] is a mini-game platform developed by Virtualware that uses the Leap Motion controller to offer physiotherapy for conditions such as Parkinson's disease and stroke. Ten Ton Raygun's Visual Touch Therapy software [9] also employs the Leap Motion controller to make therapy fun and engaging, while allowing the monitoring and assessment by therapists. The target group of these therapeutic games is mostly post-stroke patients, but also individuals with spinal cord injuries, head injuries and nerve damage. Exactly because of its accuracy and portability, the Leap Motion sensor is very promising for CTS rehabilitation.

3.1 The Leap Motion Sensor

The Leap Motion controller presents a new gesture and position tracking system with accuracy that reaches a sub-millimeter. This sensor is a small USB peripheral device which is designed to be placed on a table, facing upward. The device observes an approximately hemispherical volume to a distance of about 1 meter from it using two monochromatic IR cameras and three infrared LEDs. The LEDs generate IR light and the cameras generate almost 300 frames per second of reflected data, which is streamed through a USB cable to the host computer to be analyzed by the Leap Motion controller software to generate 3D position data for the arm, hand and fingers.

The assessment of optical 3D sensors usually depends on their stereo vision cameras, the structured light sensors and the time of flight cameras. The stereo vision cameras that are formed by two 2D optical cameras determine the depth of the environment by comparing and combining the points extracted by the two cameras. Structured light sensors inspect the distortion of a noted design to an unfamiliar area to ascertain the shape of the objects in the three dimensional space, while the so called time of flight (TOF) 3D cameras resolve the distance of the target object by measuring the time of flight of a light signal emitted by the device. Also, the devices can be divided into photonic mixer devices, which measure the phase shift between the sent infrared light beam and the reflected returned light, and laser sensors, which measure the distance to an object according to the time the reflected pulsed laser beams need to return to the sensor.

The Leap Motion controller can be classified as an optical tracking system based on stereo vision, as it emits infrared light and has two infrared cameras that calculate the position points of the target objects, such as hands, fingers or tools in the Cartesian space. The calculated position is relative to the Leap Motion controller's center point that is located at the position of the second, centered infrared emitter, as can be seen in **Error! Reference source not found.**.

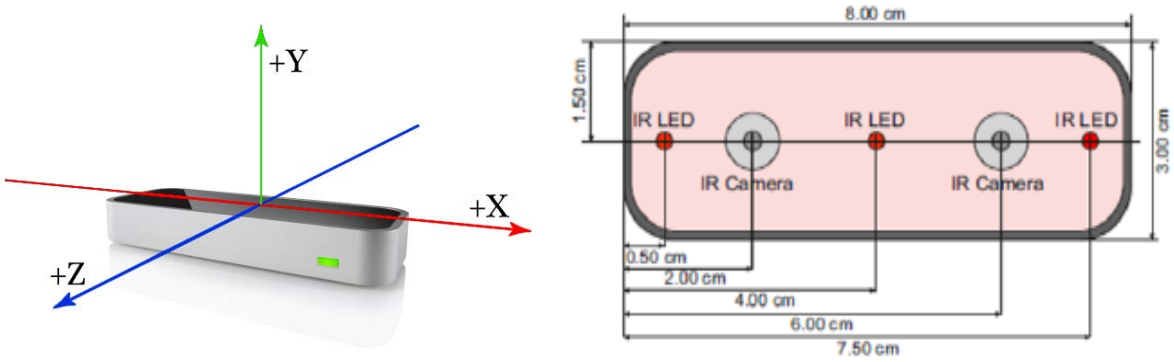


Figure 1: The Leap Motion Sensor on the left and schematic view of its internals on the right

Weichert, Bachmann, Rudak and Fisseler [23] evaluated the accuracy and repeatability of the Leap Motion controller with the use of industrial robots. The conclude that the deviation between the certain object position in the three dimensional space and that calculated by the Leap Motion sensor is less than 0.2 mm with repeatability at the level of less than 0.17 mm. Other popular sensors, like Microsoft Kinect, tested in the experiment could not achieve such accuracy results. While it is not possible to achieve the theoretical accuracy of 0.01 mm in real conditions, still the sensor offers an overall accuracy of about 0.7 mm, which is considered high precision for any hand gesture recognition user interfaces and thus makes the Leap Motion controller appropriate for physiotherapy applications.

Tracking data transmitted from the sensor to the application is in the form of snapshots called frames, which contain calculated position and rotation points and other information regarding the entities that were detected by the sensor at a certain moment of the snapshot. The physical entities that can be detected by the sensor are the hands and the fingers of the user or in some cases tools like pens or pencils.

The frame() function of the Controller class can be used to retrieve a frame object containing an entity's tracking data. This function utilizes a history parameter that specifies the number of previous frames to obtain, while the history buffer can fit the last 60 frames. Further data of the frame class can be accessed through several class functions.

The data of a detected hand of a user can include information about the position, the rotation, the direction and the motion of the hand:

- isRight, isLeft — identifies the detected hand (right or left).
- Palm Position — calculates the position of the center of the palm in millimeters relatively to the sensor's origin.
- Palm Velocity — the speed and direction of the palm in millimeters per second.
- Palm Normal — represents a vector perpendicular to the plane of the palm of the hand which points out of the palm.
- Direction — a vector pointing from the center of the palm toward the fingers.
- grabStrength, pinchStrength — variables that describe the posture of the hand.

The motion factors provide relative translation, rotation and scale differentiations in movement between different frames. In addition, the vector class includes methods that retrieve the pitch (rotation around the x-axis), yaw (rotation around the y-axis), and roll (rotation around the z-axis) rotations of the detected entity [24].

3.2 The Unity3D Game Engine

Unity3D is a popular cross-platform game development tool commonly used for the development of PC, consoles, mobile devices video games and applications. This game engine focuses on portability, targets a number of APIs and allows publishing the resultant applications in Android, Apple TV, BlackBerry 10, iOS, Linux, Nintendo 3DS, OS X, PlayStation 4, PlayStation Vita, Unity Web Player, Wii, Wii U, Windows Phone 8, Windows, Xbox 360 and Xbox One [25]. Unity's versatility includes the ability to determine the appropriate texture compression and resolution settings for each platform that it supports. In addition, the game engine provides reflection, bump, and parallax mappings, screen space ambient occlusion (SSAO), dynamic shadows using shadow maps, render-to-texture and full-screen post-processing effects. The graphics engine recognizes the most compatible modification of the shader regarding the video hardware of the system, allows alternative specifications for the shader [26] and includes an asset server and Nvidia's PhysX physics engine.

Within a simple user interface Unity3D's integrated editor shown in Figure 2 provides external editors for scripting, texturing and model editing, as well as a convenient asset and file organization system. The game engine is also highly optimized so it can support high resolution meshes without a significant frame rate decrease. In addition, assets can be imported and for textures, models and scripts all common file formats are supported.

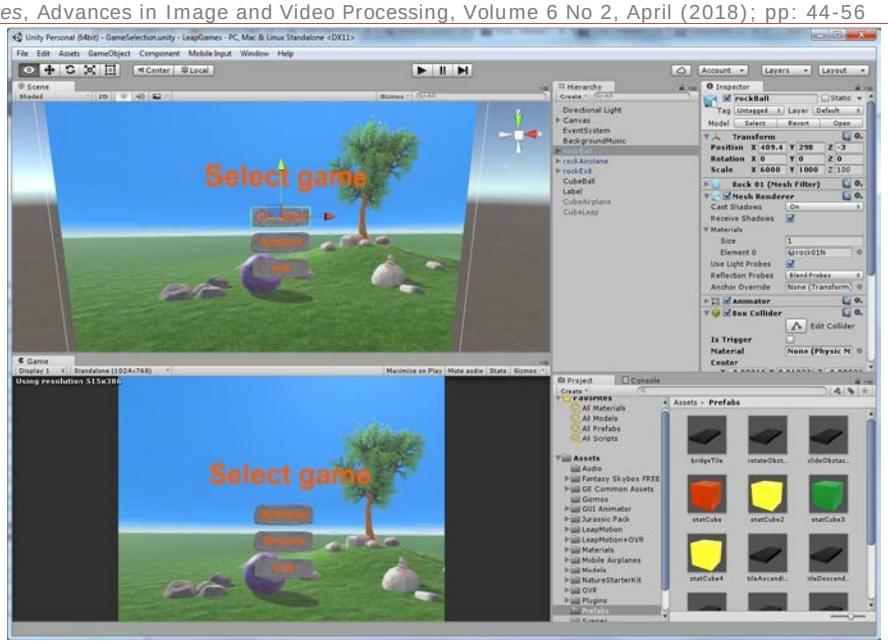


Figure 2: Developing an exergame in Unity3D for computer-aided rehabilitation of the Carpal Tunnel Syndrome.

Unity's shader system is usable and flexible; it includes shaders ranging from very basic like Diffuse or Glossy to advanced, like self-illuminated bumped specular etc. The built-in shaders can be successfully combined with any type of light. On the other hand it is possible to create custom shaders by using ShaderLab language with Cg and GLSL. The terrains of Unity are not particularly demanding in terms of hardware and perform satisfactorily on lower-end hardware, as long as the models contain a reasonable amount of polygons. The lights and the shadows available in Unity3D include real-time soft-shadows, baked light maps, halos and lens flares.

Unity3D also contains a network system that allows the development of full featured real-time multiplayer games, since it is able to access HTTP servers and establish connection with PHP websites. Regarding physics, the game engine includes a built-in physics engine named Ageia PhysX™, which supports the rigid body paradigm which reacts to the forces being applied as well as collisions and does not require scripting to function at its basic level. The game engine also provides rich libraries and thorough documentation regarding the .NET-based C# and JavaScript scripting. In addition to the native Unity editor, users may utilize the Visual Studio editor to compile scenes.

3.3 Customized Exergames for CTS

The present work augments our existing Leap Motion sensor –based CTS rehabilitation system [27] intended to offer physical therapy opportunities at home combined with the Fly-A-Plane exergame. The process benefited from a number of meetings with a physiotherapist in order to (a) glean details regarding CTS and its treatment and (b) select a number of physiotherapy exercises to be performed by the player (patient) to control game progress. That initial exercise set **Error! Reference source not found.Error! Reference source not found.**contains ten distinct movements of the wrist and the fingers, including different rotations and postures of the wrist and finger joints. Out of that initial set, the four exercises that were selected to implement to control the Fly-A-Plane game appear in Figure 3. The criterion for the given selection stems from the need that, since the movements of the exercises must correspond to the movements of the avatars in the game environment, the user should naturally perform these exercises without confusion.

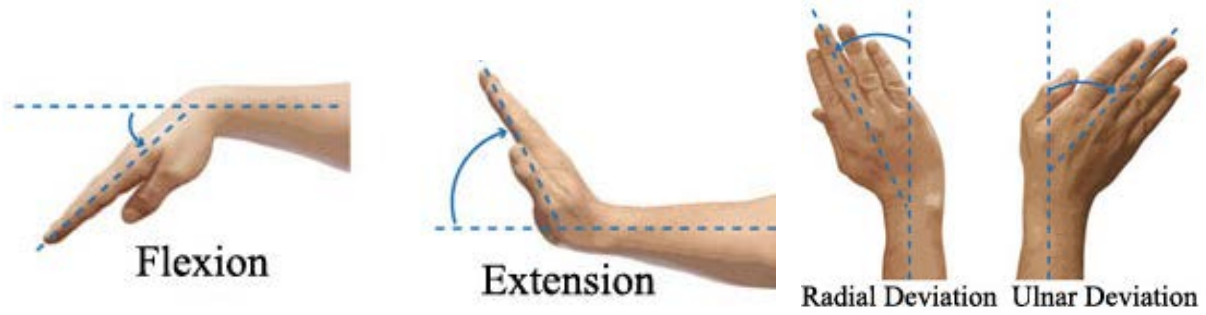


Figure 3: – The four exercises selected for implementation in the Fly-A-Plane exergame.

The goal of the Fly-A-Plane exergame is to guide an airplane through a number of hoops in the sky (game scene). If the Leap Motion sensor detects one of the hand movements listed in Figure 3, the airplane is directed to perform a pre-programmed movement corresponding to that particular detection event. Therefore, the physical movement of the player's arm corresponds very naturally to the game avatar (airplane) movement: positive pitch rotation for a wrist extension, negative pitch rotation for wrist flexion, positive yaw rotation for ulnar deviation and negative yaw rotation for radial deviation as shown in Figure 4 and detailed in Table.

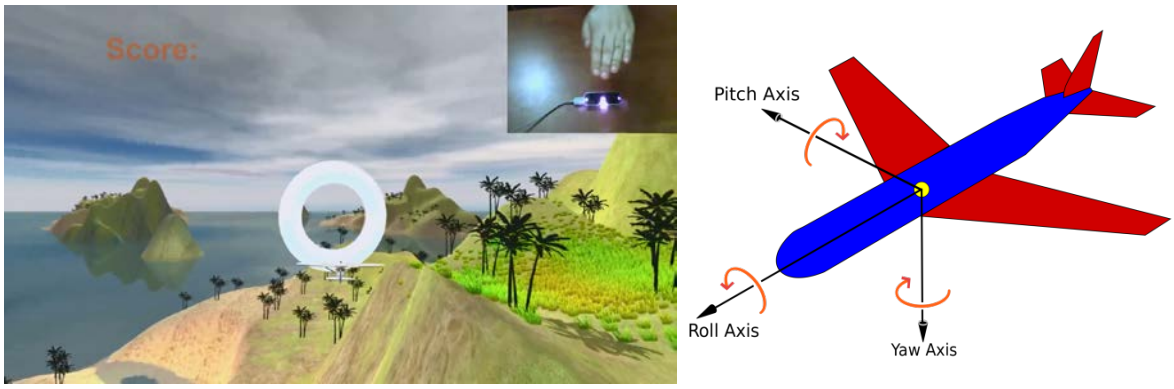


Figure 4: – On the left appears a selected main scene of the Fly-A-Plane exergame. The insert in the top right shows the player's hand above the Leap sensor guiding the plane. The principal axes of rotation of the plane appear in the right figure.

Table 1. Correspondence between physical and game avatar (airplane) movements in the Fly-A-Plane exergame

Physical therapy exercise movement	Game avatar movement
Wrist extension (upward rotation)	Positive pitch angle rotation
Wrist flexion (downward rotation)	Negative pitch angle rotation
Wrist ulnar deviation (rotation to the right)	Positive yaw angle rotation
Wrist radial deviation (rotation to the left)	Negative yaw angle rotation

The scene of the Fly-A-Plane game is constantly generated as long as the player guides the airplane in any direction. The scene is created by generating a 3x3 grid made of the initial terrain, containing exotic islands and sea environment and is repeated in a specific direction from the game starting point, whenever the airplane reaches that area. The need for generating the grid is coded in the Update() method so that it is checked in every frame, and its position is determined in the UpdateTerrainPositionsAndNeighbors() method listed in Figure.

```
void Start ()
{
Terrain linkedTerrain=gameObject.GetComponent<Terrain>();
_terrainGrid[0,0]=
Terrain.CreateTerrainGameObject(linkedTerrain.terrainData).GetComponent<Terrain>();
_terrainGrid[0,1]=
Terrain.CreateTerrainGameObject(linkedTerrain.terrainData).GetComponent<Terrain>();
_terrainGrid[0,2]=
Terrain.CreateTerrainGameObject(linkedTerrain.terrainData).GetComponent<Terrain>();
_terrainGrid[1,0]=
Terrain.CreateTerrainGameObject(linkedTerrain.terrainData).GetComponent<Terrain>();
_terrainGrid[1,1]=linkedTerrain;
_terrainGrid[1,2]=
Terrain.CreateTerrainGameObject(linkedTerrain.terrainData).GetComponent<Terrain>();
_terrainGrid[2,0]=
Terrain.CreateTerrainGameObject(linkedTerrain.terrainData).GetComponent<Terrain>();
_terrainGrid[2,1]=
Terrain.CreateTerrainGameObject(linkedTerrain.terrainData).GetComponent<Terrain>();
_terrainGrid[2,2]=
Terrain.CreateTerrainGameObject(linkedTerrain.terrainData).GetComponent<Terrain>();
UpdateTerrainPositionsAndNeighbors();
}
```

Figure 5: – Sample code responsible for the creation of a 3x3 grid repeating the initial terrain object.

```
controller=new Controller();

Frame frame=controller.Frame();
yaw=frame.Hands[0].Direction.Yaw;
pitch=frame.Hands[0].Direction.Pitch;

if(frame.Hands.IsEmpty){airplane.transform.Translate(0f,0f,-0.01f* speed);}
if(!frame.Hands.IsEmpty && pitch >minPitch+0.3){
    Vector3 movementUp=new Vector3 (0f,-0.1f,0f);
airplane.transform.Translate(movementUp);
transform.eulerAngles=new Vector3
(Mathf.LerpAngle(transform.eulerAngles.x,30.0f,Time.deltaTime),
transform.eulerAngles.y,Mathf.LerpAngle(transform.eulerAngles.z,180.0f,Time.deltaTime));
…
```

Figure 6: – Sample code implementing the airplane movement controller.

Since the main goal of this game is to fly the airplane through the rings in the game environment, it is exactly the sequence of positions of the rings that is crucial for the correct execution of the physical therapy exercises. For that reason the rings are being instantiated dynamically according to rules implemented according to the directions / requirements stated by the physiotherapist for the particular patient. For example, if the user has just performed an ulnar deviation to turn the airplane to the right to pass through a ring and the script instructs for a radial deviation to be performed next, the new ring must appear to the left of the plane's new direction. Similarly, when the next ring is situated in front of the airplane and at a higher altitude, the user is forced to perform an extension of the wrist to make the airplane to climb to reach that ring, whereas a ring positioned at a lower altitude requires a flexion of the wrist. In this way, the patient naturally executes a prescribed exercise script in a meaningful and engaging game environment. Essentially, this is the job of the airplane movement controller detailed in Figure.

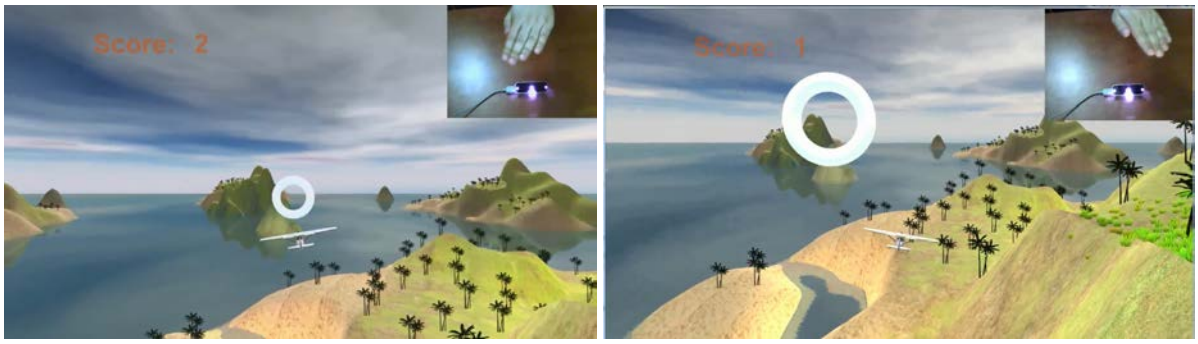


Figure 7: – Left panel: To turn the airplane left requires a radial deviation of the right hand or an ulnar deviation of the left hand. Right panel: To turn the airplane to the right requires an ulnar deviation of the right hand or a radial deviation of the left hand

In the left panel in Figure the next ring to be collected is located in front and to the left of the airplane, forcing the user to rotate the plane to the left. This can be accomplished using a radial deviation of the right hand or an ulnar deviation of the left hand. By comparison, the right rotation of the plane required in the right panel in Figure requires either an ulnar deviation of the right hand or a radial deviation of the left hand. (Each panel shows a mirrored image of the hand in the upper right corner of each panel's game scene shows a mirrored image of the hand above the Leap Motion controller.)

Unlike the first game developed for our CTS-rehabilitation platform (see Roll-A-Ball game in [27]) where wrist flexion and extension were employed as "go" and "stop" actions on the rolling ball, respectively, in the Fly-A-Plane exergame the airplane must be allowed to gain or lose altitude, meaning that exercise patterns detected by the Leap Motion sensor are being mapped onto game actions based on the specific abilities of the controlled game objects. The airplane moves forward constantly, but with a slow speed, so that the game is more challenging, while the user still has enough time to perform the exercise at a normal pace. The insert in the left and right panels in Figure shows the wrist extension / flexion exercise being performed in order to move the airplane up / down, respectively.

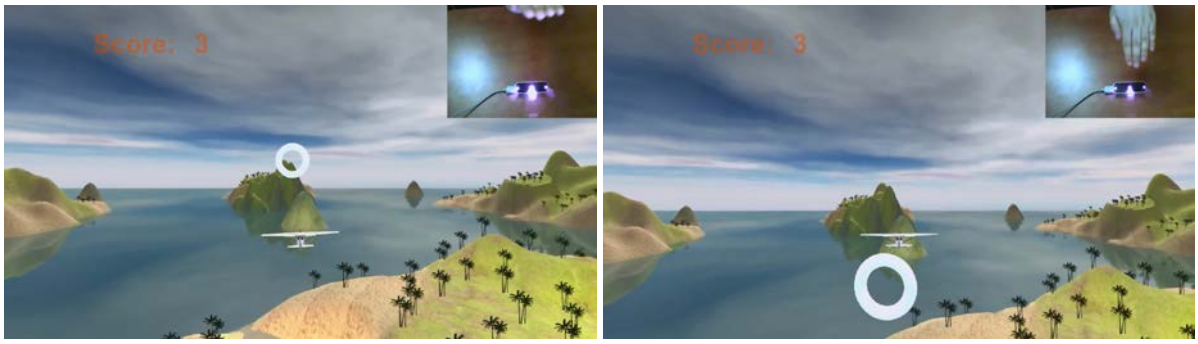


Figure 8: – Left panel: Execution of a wrist extension to cause the airplane to gain altitude; Right panel: Execution of a wrist flexion to cause the airplane to lose altitude..

It is worth noting in Figure 3 that the angular range of normal wrist rotation is not the same in both ulnar and radial deviation in the sense that the range of motion for the ulnar deviation of the right hand or the radial deviation of the left hand is considerably more wide than the possible range of motion for the ulnar deviation of the left hand or the radial deviation of the right hand. Accordingly, the detection limits for the Leap motion sensor were set higher in the former than in the latter case. Recognizing that CTS severity also differs among patients, additional customization is possible and each exercise is parametrized to be adapted to the needs of the individual patient. For demo purposes the system presently allows users to select a predefined difficulty level (easy, medium and hard) for each game they play (shown in the left panel in Figure). A harder difficulty level yields a higher speed of movement

for the avatar and implies more extreme wrist rotation angles. For example, if the user chooses to play in easy mode, the airplane will be moving forward slower and it will require smaller wrist rotations to turn the plane up or down or change its altitude. This can be a level appropriate for a beginner or a more advanced case of CTS.

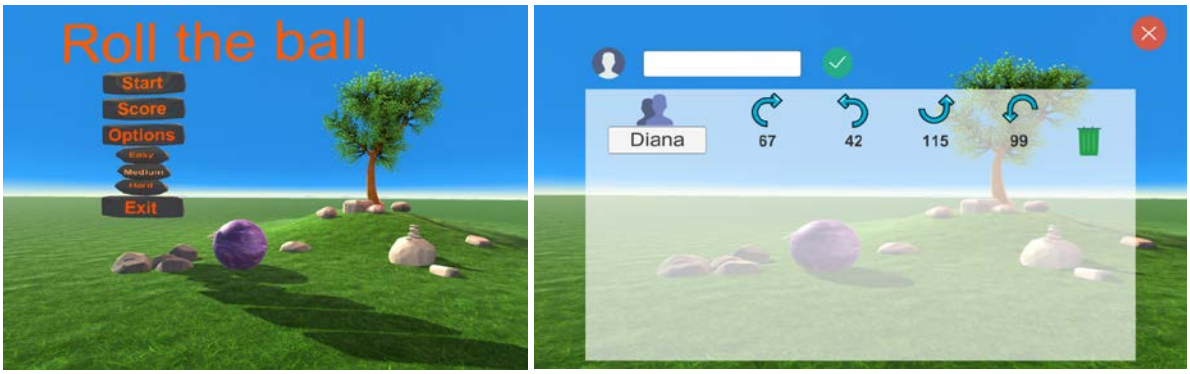


Figure 9: – Left: Top level showing the difficulty levels; Right: Detailed user performance per game cycle..

Finally, the system stores high scores for every hand rotation (right, left, up and down) for every user profile as shown in the right panel in Figure and sends relevant data with JSON (JavaScript Object Notation) to a server, which displays the score graph for every user by the date of the game play.

4 Discussion and Future Work

In this work we enrich our Carpal Tunnel Syndrome (CTS) rehabilitation platform [27] with a Fly-A-Plane exergame aimed at alleviating symptoms of the syndrome for suffering patients. The game has been developed Unity3D and capitalizes on the ability of the Leap Motion sensor to accurately track wrist-hand movement and detect certain pre-programmed rehabilitation movements. At first glance, the use of the Leap Motion sensor may appear limited for rehabilitation, because it tracks only the lower arm, hand and fingers. However this is sufficient for serious games tailored to CTS patients.

One possible direction for future work is to examine whether and to what degree of usefulness the Leap Motion sensor can be paired with Microsoft's Kinect to provide the required precision in the detection of hand gestures like pronation/supination and flexion/extension of metacarpophalangeal and the proximal interphalangeal joints of the hand, while Kinect provides larger scale body tracking.

Finally, technology-enabled rehabilitation can offer measurable results as games can be designed to collect performance data and use these to calculate and store appropriate performance metrics. This possibility should allow therapists to assess long term progress and to mitigate a sense of accomplishment to the patients which can be very beneficial during treatment.

ACKNOWLEDGEMENTS

The authors thank physiotherapist Mr. George Stratakis for his input in gleaning exercises that are feasible to model using the Leap Motion Sensor.

REFERENCES

[1] M. Chammas, J. Boretto, L. M. Burmann, R. M. Ramos, F. C. dos Santos Neto, and J. B. Silva, "Carpal tunnel syndrome – Part I (anatomy, physiology, etiology and diagnosis)," *Rev. Bras. Ortop. (English Ed.*, vol. 49, no. 5, pp. 429–436, 2014.

[2] A. Alexander, "Carpal tunnel syndrome - Information for Physiotherapists," *Oxford University Hospitals NHS Trust*. 2014.

[3] T. Gregory, D. Chamblin, J. Firestone, A. Friadman, H. Chirstopher, P. Douglas, M. Kliot, R. Lawrence, E. Thomas, B. Nicholas, D. Michael, T. Kjerulf, K. O'Bara, S. Carlson, and J. G, "Work-Related Carpal Tunnel Syndrome Diagnosis and Treatment Guideline," *Washingtonworkingsolutions.Net*. Washington State Department of Labor & Industries, pp. 1–18, 2017.

[4] M. Abdur Rahman, "Multisensor Serious Game-Based Therapy Environment for Hemiplegic Patients," *Int. J. Distrib. Sens. Networks*, vol. 11, no. 6, p. 910482, Jun. 2015.

[5] G. Alankus, R. Proffitt, C. Kelleher, and J. Engsberg, "Stroke Therapy through Motion-Based Games," *ACM Trans. Access. Comput.*, vol. 4, no. 1, pp. 1–35, Nov. 2011.

[6] A. Neil, S. Ens, R. Pelletier, T. Jarus, and D. Rand, "Sony PlayStation EyeToy elicits higher levels of movement than the Nintendo Wii: Implications for stroke rehabilitation," *Eur. J. Phys. Rehabil. Med.*, vol. 49, no. 1, pp. 13–21, 2013.

[7] A. M. D. C. Souza, M. a Gadelha, E. a G. Coutinho, and S. R. Santos, "A video-tracking based serious game for motor rehabilitation of post-stroke hand impairment," *SBC J. 3D Interact. Syst.*, vol. 3, no. 2, pp. 37–46, 2012.

[8] A. De Graaf, "Gaming Motion tracking technologies for rehabilitation," *Study Tour Pixel 2010*, pp. 1–6, 2010.

[9] A. Gupta, "AXLR8R Spotlight: Making Physical Therapy Fun with Ten Ton Raygun," 2017. [Online]. Available: http://blog.leapmotion.com/axlr8r-spotlight-making-physical-therapy-fun-with-ten-ton-raygun/.

[10] "What if Physical Rehabilitation Were as Easy as Playing a Video Game? - Not Impossible." [Online]. Available: http://www.notimpossiblenow.com/lives/what-if-physical-rehabilitation-were-as-easy-as-playing-a-video-game. [Accessed: 05-Mar-2016].

[11] D. L. McLellan and M. Swash, "Longitudinal sliding of the median nerve during movements of the upper limb," *J. Neurol. Neurosurg. Psychiatry*, vol. 39, no. 6, pp. 566–570, 1976.

[12] R. M. Szabo, B. K. Bay, N. A. Sharkey, and C. Gaut, "Median nerve displacement through the carpal canal," *J. Hand Surg. Am.*, vol. 19, no. 6, pp. 901–906, Nov. 1994.

[13] S. L. Michlovitz, "Conservative Interventions for Carpal Tunnel Syndrome," *J. Orthop. Sport. Phys. Ther.*, vol. 34, no. 10, pp. 589–600, Oct. 2004.

[14] L. M. Rozmaryn, S. Dovelle, E. R. Rothman, K. Gorman, K. M. Olvey, and J. J. Bartko, "Nerve and tendon gliding exercises and the conservative management of carpal tunnel syndrome," *J. Hand Ther.*, vol. 11, no. 3, pp. 171–179, Jul. 1998.

[15] K. Lohse, N. Shirzad, A. Verster, N. Hodges, and H. F. M. Van der Loos, "Video Games and Rehabilitation," *J. Neurol. Phys. Ther.*, vol. 37, no. 4, pp. 166–175, Dec. 2013.

[16] F. Anderson, M. Annett, and W. F. Bischof, "Lean on Wii: Physical rehabilitation with virtual reality Wii peripherals," *Stud. Health Technol. Inform.*, vol. 154, pp. 229–234, 2010.

[17] A. Y. Wang, "Games for Physical Therapy," *SIGCHI Conf.*, pp. 4–8, 2012.

[18] M. E. Kho, A. Damluji, J. M. Zanni, and D. M. Needham, "Feasibility and observed safety of interactive video games for physical rehabilitation in the intensive care unit: a case series," *J. Crit. Care*, vol. 27, no. 2, p. 219.e1-219.e6, Apr. 2012.

[19] A. de Mauro, "Virtual Reality Based Rehabilitation and Game Technology," in *1st International Workshop*

on Engineering Interactive Computing Systems for Medicine and Health Care (EICSMed 2011), co-located with the ACM SIGCHI Symposium on Engineering Interactive Computing Systems (EICS 2011), 2011, pp. 48–52.

[20] M. Khademi, H. Mousavi Hondori, A. McKenzie, L. Dodakian, C. V. Lopes, and S. C. Cramer, "Free-hand interaction with leap motion controller for stroke rehabilitation," in *Proceedings of the extended abstracts of the 32nd annual ACM conference on Human factors in computing systems - CHI EA '14*, 2014, pp. 1663–1668.

[21] L. Geurts, V. Vanden Abeele, J. Husson, F. Windey, M. Van Overveldt, J.-H. Annema, and S. Desmet, "Digital games for physical therapy: fulfilling the need for calibration and adaptation," in *Proceedings of the fifth international conference on Tangible, embedded, and embodied interaction - TEI '11*, 2011, vol. 53, p. 117.

[22] "Virtualrehab." [Online]. Available: http://www.virtualrehab.info/.

[23] F. Weichert, D. Bachmann, B. Rudak, and D. Fisseler, "Analysis of the Accuracy and Robustness of the Leap Motion Controller," *Sensors*, vol. 13, no. 5, pp. 6380–6393, May 2013.

[24] "Leap Motion API Reference." [Online]. Available: https://developer.leapmotion.com/documentation/csharp/api/Leap_Classes.html?proglang=csharp. [Accessed: 05-Mar-2016].

[25] J. Bae and A. Kim, "Design and Development of Unity3D Game Engine-Based Smart SNG (Social Network Game)," *Int. J. Multimed. Ubiquitous Eng.*, vol. 9, no. 8, pp. 261–266, Aug. 2014.

[26] S. Wang, Z. Mao, C. Zeng, H. Gong, S. Li, and B. Chen, "A new method of virtual reality based on Unity3D," *2010 18th Int. Conf. Geoinformatics*, pp. 1–5, 2010.

[27] I. Pachoulakis and D. Tsilidi, "Technology-assisted Carpal Tunnel Syndrome Rehabilitation using serious games: the Roller Ball example," *Adv. Image Video Process.*, vol. 4, no. 4, Aug. 2016.

AIVP ADVANCES IN IMAGE AND VIDEO PROCESSING
SOCIETY FOR SCIENCE AND EDUCATION UNITED KINGDOM
ISSN 2054-7412

Carpal Tunnel Syndrome: Causes, Prevention, Rehabilitation and Computer-Aided, Game-Based Physiotherapy

[1]Diana Tsilidi, [2]Ioannis Pachoulakis, [3]Anastasia Analyti
[1,2]Department of Informatics Engineering, Technological Educational Institute of Crete Heraklion, Crete, Greece;
[3]Institute of Computer Science, Foundation for Research and Technology – Hellas (FORTH) Vassilika Vouton, Heraklion, Crete, Greece;
dtsilidi@gmail.com; ip@ie.teicrete.gr; analyti@ics.forth.gr

ABSTRACT

Carpal Tunnel Syndrome (CTS) is caused by improper computer posture and usage of the arms and hands. It expresses a conglomerate of symptoms caused by the compression of the median nerve as it passes through the carpal tunnel of the wrist. These include pain, numbness and tingling in affected parts of the hand. CTS treatment may take the form of surgery and/or physiotherapy. This paper revisits the causes of CTS and the common prevention tactics and focuses on the emerging role of rehabilitation via computer-aided, game-based physiotherapy which can augment a more traditional exercise curriculum.

Keywords: Carpal tunnel syndrome (CTS); Computer-aided rehabilitation; Game-based Rehabilitation.

1 Introduction

The Carpal Tunnel Syndrome (CTS) is a medical condition in which the median nerve is compressed as it travels through the wrist's carpal tunnel causing pain, numbness and tingling in parts of the hand that receive sensation from the median nerve [1]. It has also been associated with conditions such as obesity, hypothyroidism, arthritis, diabetes and trauma. The syndrome was at first described by Sir James Paget, after a case of medial nerve compression at the wrist following a fracture of the distalradius. The principles of CTS were established in 1950, when George Phalen reported for the very first time this relatively unknown condition at the 99th Annual Meeting of the American Medical Association. To comprehend CTS, one must first get familiarized with the relevant wrist anatomy and the tunnel that is shaped by the carpal bones as shown in **Error! Reference source not found.**. The wrist has an incurved contour that consists of eight bones, forming the joint, on the surface of the flexor, the muscle which bends the limb by its contraction, and is covered by the flexor retinaculum or transverse carpal ligament, a fibrous band on the palmar side of the hand [2].

In effect, the carpal tunnel is confined by the bony carpus and the transverse carpal ligament (flexor retinaculum). The latter connected to the tubercle of the scaphoid carpal bone, the ridge of the trapezium carpal bone and the ulnar side of the hook (small projection of the bone) of the hamate and pisiform carpal bones. The flexor muscles of the fingers pass through the carpal tunnel, while the median nerve is located under the transverse carpal ligament [2].

In conclusion, the wrist consists of small bones, the carpal bones, which form the carpal tunnel. The median nerve, which controls the sensations in the palm and fingers, as well as the tendons pass through this tunnel from the forearm into the hand. The pressure on the median nerve can be

DOI: 10.14738/aivp.62.4459
Publication Date: 29th April 2018
URL: http://dx.doi.org/10.14738/aivp.62.4459

sometimes caused by swelling and irritation of the tendons, thus producing the symptoms of the CTS. The affected hand is usually the dominant hand of the patient, although it also quite possible to observe symptoms of CTS in both hands [2].

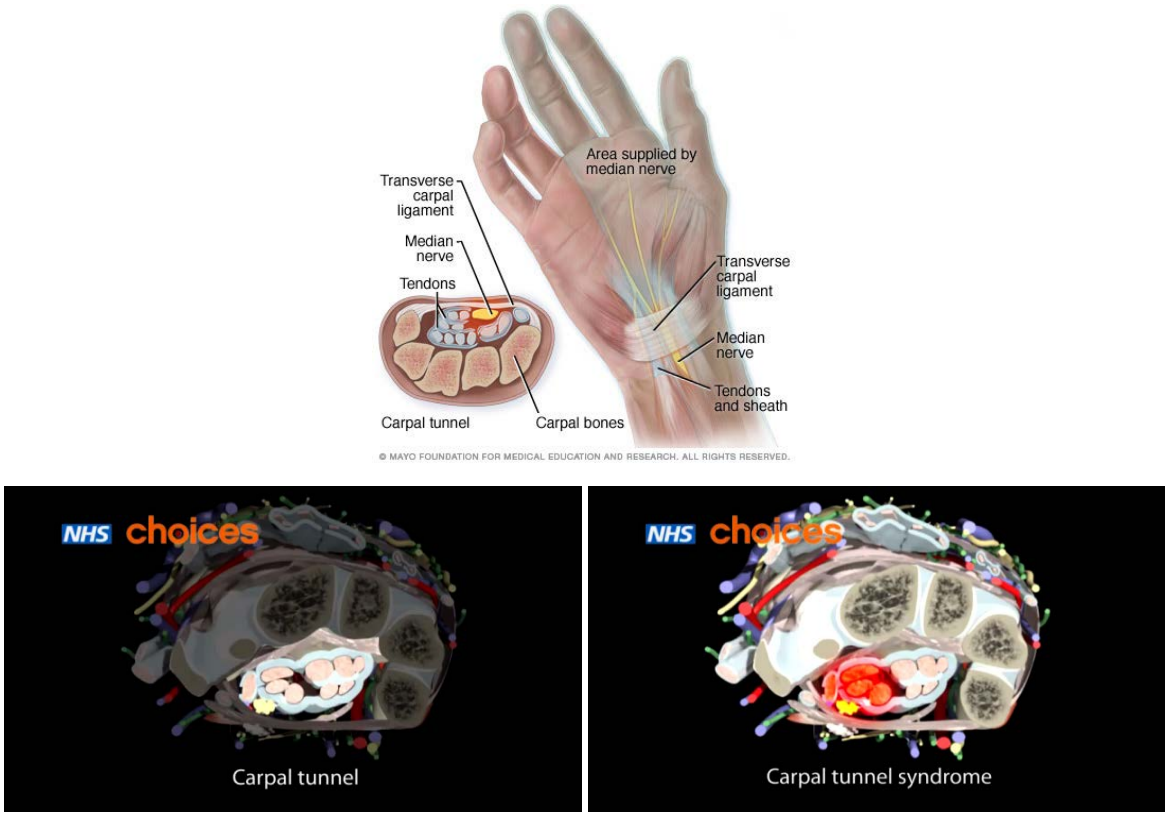


Figure 1: Top: Wrist anatomy (adopted from [3]); Bottom Left: Physiological ligaments in the carpal tunnel; Bottom right: Inflamed ligaments in red in the carpal tunnel (both images adopted from [4])

Treatment of CTS aims at restoring normal hand functionality to allow patients to resume their usual daily activities as well as prevent further/additional nerve damage and muscle loss in the fingers and hands. Mainstream treatment options include wrist splinting and immobilization, administration of medicines such as corticosteroids and anti-inflammatory drugs that relieve pain and reduce inflammation, physical therapy and, finally, surgery as a last resort measure, when other treatments fail.

2 Origins of the Carpal Tunnel Syndrome

The compression of the median nerve occurs at the carpal tunnel as the result of an incompatibility between the relative size of the canal and the volume of its content, the space occupied by the transit tendons and the median nerve. Studies, including the measurement of the inter-carpal canal pressure of CTS patients, have shown significantly increased mean pressure in the carpal canal. The most significant mean carpal canal pressure increment was observed during the wrist extension and slightly less, when wrist flexion was performed, while an immediate pressure decrease was observed during carpal tunnel release. Therefore, the syndrome is usually caused by several different conditions. The space inside the carpal canal can be reduced by processes including tenosynovitis of the flexor tendons, fracture or dislocation of the carpus and carpometacarpal joints and Colles' fracture. These conditions may also cause post-traumatic scarring or fibrosis inside the carpal tunnel. Rheumatoid arthritis, amyloid deposition, granoulomatous infection and gout are some of the inflammatory processes that contribute to reduced volume inside the carpal tunnel and produce a proliferative tenosynovitis with

hyperplastic synovium [2]. Other causes for the decreased space within the carpal canal can be tumors of the median nerve, disorders like acromegaly or hypothyroidism, pregnancy, diabetes mellitus, and lupus erythematosus. Some of these conditions may also increment the extracapsular fluid retention and produce swelling of the soft tissue.

It has been observed that women are three times more likely to have CTS than men. However, the reason for this increased frequency in women population is not known yet. Some of the theories mention that it may be caused by the difference in the size of the wrist bones, which are naturally smaller in most women. This means that the space that is formed by the carpal bones, through which the nerves and tendons pass, is tighter. Other factors that are considered by the scientists are the genetic links that make it more likely for women to have musculoskeletal injuries and the strong hormonal changes.

Certain health problems can also contribute to the onset of CTS, although the exact origins may remain unknown. Factors that raise the chances of developing CTS include genetic predisposition, wrist injury causing swelling and pressure on the nerve or repetitive movements of the hands (computer users, assembly line workers, carpenters, musicians, athletes etc.). Additional women-specific factors contributing to the development of CTS relate to hormonal changes, pregnancy and menopause. In these cases the wrist structure may become enlarged and add pressure on the nerve, but in many cases the symptoms fade after the condition passes. Another rare cause of CTS can be breast cancer, since often after mastectomy the build-up of fluids raises beyond the ability of lymph system to drain it, causing pain and swelling of the arm and pressuring the nerve.

In detail, the most significant and common medical conditions that lead to CTS can be classified into the groups mentioned below:

- Swelling-tenosynovitis of the flexor tendons of the hand due to extensive hand use (often observed to farmers, builders, hairdressers, drivers or wheeled machine operators).
- Swelling of the tendon due to hormonal disorders such as disorders of the thyroid gland and ovarian (menopause or pregnancy).
- All of the above conditions, causing swelling of the sheaths and tendons within the carpal tunnel leading to an increase of pressure within. Initially the pressure affects small vessels of perineural and then the whole nerve that results to conduction disorder. The nerve swells and presents dotted macroscopic bleeding. In chronic cases of this neglected condition of the wrist the medium nerve presents intraneural fibrosis and permanent functional impairment.
- Rarely the syndrome is caused by diabetes or a congenital narrowing of the carpal tunnel.
- In many cases CTS is a result of incorrect positioning of the wrist when typing on a computer. That is why the newer keyboards have wrist support extensions. The disease has also been called journalistic and "typists disease" or "disease of the computer operator."
- In recent years it has been discovered that the intake of the drug Tamoxifen and Arimidex may also lead to CTS.
- In rare cases, CTS coexists with median nerve pressure on the forearm, by the round pronator muscle.
- They have been rare but existing cases of CTS caused by tumors and CTS amyloidosis of histological lesions.

According to [5], the main causes and contributing factors to CTS are:

- Aberrant Anatomy: anomalous flexor tendons, congenitally small carpal canal, ganglionic cysts, lipoma, proximal lumbrical muscle insertion, thrombosed artery.

- Infections: lyme disease, mycobacterial infection, septic arthritis.
- Inflammatory conditions such as connective tissue disease, gout or pseudo-gout, rheumatoid arthritis and nonspecific flexor tenosynovitis (the most common cause of CTS).
- Metabolic conditions such as diabetes, hypothyroidism / hyperthyroidism, acromegaly, amyloidosis.
- Increased canal volume: congestive heart failure, edema, obesity, pregnancy.

It has also been observed that smokers with CTS have worse symptoms and usually recover more slowly than non-smokers.

3 Symptoms of Carpal Tunnel Syndrome

Proper CTS diagnosis entails the observation of the appropriate symptoms in conjunction with abnormal nerve conduction test. The telltale symptoms of CTS are burning pain in the volar aspects of both hands, numbness in the fingers and tingling, which often become worse after work and at night and awakens patients from their sleep. Nighttime symptoms are prominent in 50-70% of the patient population. Shaking the wrist down may relieve the symptoms. Symptoms are commonly localized to the thumb and the first two or three fingers, often accompanied by burning pain in the palm side of the wrist (although this sensation of pain, numbness and tingling may also expand to the forearm, elbow, and especially the shoulder in the form of aching pain. In severe cases, patients may even lose the hand dexterity because of decreased grip strength and thenar muscle atrophy. Ulnar neuropathy or C-8 radiculopathy should be considered in cases where the nerve symptoms are observed only in the fourth and fifth fingers, whereas proximal symptoms, especially tingling in the radial hand combined with elbow side pain may point towards a possible C-6 radiculopathy.

Signs and symptoms observed during the physical examinations may often be negligible or confusing. Some of the diagnosis methods, like Hoffmann-Tinel's sign, which is described as paresthesias expanding in the median nerve distribution with tapping on the wrist or over the median nerve, and Phalen's sign, which involves paresthesia radiating in a median nerve distribution, but within 60 seconds of sustained flexion of the wrist, are frequently described, though are not specific enough for the diagnosis of CTS. The presence of these signs reinforces the existence of other specific neurologic symptoms. In addition, non-specific symptoms, like pain without numbness, tingling and burning cannot be considered as diagnostic of CTS by themselves.

More severe cases of CTS may involve additional symptoms, including decreased sensation to pin or light touch in the first three digits or weakness or atrophy of the muscles of the thenar eminence, especially the abductor pollicis brevis. Thenar weakness or atrophy may reveal more acute or advanced nerve injury, in contrast with Hoffmann-Tinel's signor Phalen's sign, and in this case possibly creates the need for more aggressive treatment.

It is important to exhaust every possible effort to objectively verify the diagnosis of CTS before considering surgery. It has been showed that patients who have undergone carpal tunnel surgery with normal or near normal pre-surgical nerve conduction test results have worse outcomes than those with electro-diagnostic evidence of median nerve entrapment across the carpal tunnel, even though some evidence disagree. A rare method of therapeutic and diagnostic challenge test can be a steroid injection performed into the carpal canal. If the patient experiences improvement in symptoms for weeks or months after the injection but then relapses, surgery may be considered as a solution for carpal tunnel release. Further diagnostic evaluation can be performed by an appropriate specialist for the patients with a negative response [2].

Therapists should also investigate other clinical problems that may possibly relate to work exposure [6], such as tendonitis, in cases where CTS is not diagnosed by clinical criteria and nerve conduction testing. In addition, the patient should be referred to an appropriate specialist to eliminate the possibility of having other neurologic causes of tingling in the hands.

3.1 Associated conditions

Because of the functional and structural changes caused by CTS or other forms of polyneuropathy that lead to the syndrome, the median nerve becomes very sensitive to different compressive conditions. One of the associated conditions is hereditary neuropathy with hypersensitivity to pressure, which is a hereditary sensory motor form of neuropathy that is central and repetitive. The symptoms of the particular neuropathy, which include paralysis and paresthesias of the nerve trunk, usually appear after the age of 20. The symptoms, which usually occur after an injury or extended compression, often deteriorate and become more frequent. The condition is characterized by a disorder affecting the myelin of peripheral nerve fibers with stretching of the distal motor latency, which causes focal thickening of myelin in some zones of the nerve.

Another associated condition is the double constriction syndrome, which is based on the evidence that direct compression on the path of a nerve makes it more vulnerable than if the compression was located more peripherally. An appropriate clinical examination in addition with an electrophysiological study can determine whether the compression location is direct or peripheral and thus the treatment can be applied to the main compression location [5].

4 Diagnosis

Diagnosis CTS for patients reporting abnormal sensation like tingling or numbness of the hand and fingers, should include several steps [1], as described below. At first, the appropriate specialist should have the initial discussion with the patient regarding the symptoms and consider possible pathological conditions associated with the described symptoms. At this phase of the diagnosis it is important to conduct challenge tests, to be able to determine the etiology of the condition.

Some of the most common challenge tests are the following:

- Tinel sign - this test shows whether the patient feels tingling and numbness of the hands and fingers, when the specialist manually taps the palmar side of the wrist at the level of the median nerve. In that case the test is positive.
- Phalen sign – during this test the patient must perform maximum active flexion of the wrist for 1 minute, while having the elbow extended. If paresthesia appears in the area of the median nerve, the test is considered positive.
- Paley and McMurphy test – the particular test shows the presence or absence of paresthesia or pain triggered by the manual pressure about 1 centimeter from the median nerve.
- Compression test with wrist flexed – during this test the specialist uses his two fingers to apply pressure on the median region of the carpal tunnel, while the wrist is being flexed at 60 degrees, the elbow is extended and the forearm is supinated. In case of the appearance of paresthesia in the area of the median nerve, the test is positive.

Also, the presence of acroparesthesia at night, which includes tingling, numbness, swelling or pain, perceptible in the first three fingers, is considered one of the most sensitive tests.

The next step is to perform the Weber test, which evaluates the severity of the nerve compression, and analyze the strength of the thenar muscles innervated by the median nerve. The specialist then has to decide if additional examinations, such as nerve conduction velocity test (**Error! Reference source not**

found.), are necessary. It is especially important to perform the nerve conduction velocity test, in cases where carpal tunnel decompression surgery is considered [7].

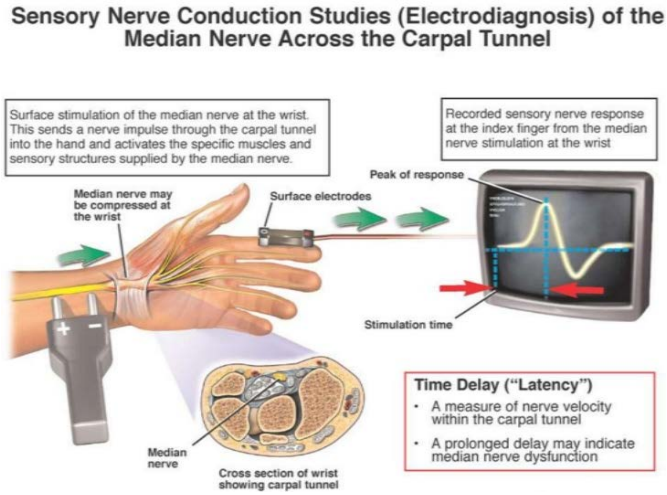


Figure 2: Nerve conduction study (adopted from [8])

The clinical diagnosis of CTS is validated by the nerve conduction study with high degree of reliability, when there is evidence of slowing of sensory and motor fibers of the median nerve across the carpal tunnel. In case all the symptoms of the patient indicate a positive clinical picture of CTS, while the nerve conduction test results are negative, the physician should consider other clinical diagnoses like tendonitis, pronator syndrome or cervical radiculopathy. The final step of the diagnosis procedure is to propose an appropriate treatment adjusted to the location, the cause and the severity of the condition.

5 Treatment

CTS patients can benefit from condition-specific guidelines that help make the treatment more effective and relieve the symptoms quickly. One of the most important guidelines is to avoid repetitive wrist and hand motions if possible and stop using vibratory tools that cause deterioration of the symptoms, such as jackhammers, floor sanders etc. Also, patients have to consider changing their work conditions, like changing the position of the wrists while working on the computer or start using wrist supports.

5.1.1 Prevention

Common musculoskeletal work-related injuries lead to CTS in improper workplace conditions. In order to prevent such injuries, the workspace and the equipment that used have to be at the right height and distance for the hands, so that the work is performed without straining the wrists. Computer users must maintain correct posture and ensure that keyboard position allows the wrist to lay comfortably without bending. Regular breaks also help lower the risk of swelling.

Other ways to prevent CTS is to perform tasks that use different muscle movements and try to do the hand and wrist motions without unnecessary tension. Exercise, like flexing and bending the wrists, after tasks demanding repetitive movements can reduce the negative effects of these tasks. Keeping the muscles warm makes them less likely to get hurt, so it is important to keep the hands warm at work, sometimes by wearing gloves.

5.1.2 Splints

Splints are being used to support and brace the wrist, keeping it in a neutral position so that the nerves and tendons can recover. They are known to reduce repetitive flexion and rotation of the wrist and relieve mild swelling of the soft tissue or tenosynovitis. Splinting the wrist is usually more effective when it is performed within three months of the manifestation of the symptoms and when it is adjusted to the symptoms and the preferences of the patient. In addition, wearing the splints full time provides more effective improvement of the symptoms and electrophysiologic measurements.

5.1.3 Surgery

Surgery is the last resort for CTS patients and should only be considered in severe cases, when conservative measures do not alleviate the symptoms and other types of treatment do not have the expected result within at least six months of therapy [9]. The outpatient procedure for surgery includes a regional anesthesia that causes numbness to the wrist and hand area and a small incision in the palm that facilitates a division of the transverse carpal ligament and its overlying structures and enlarges the carpal tunnel. After the procedure the wrist needs to be splinted for three to four weeks. Possible surgery-related complications include injury to the palmar cutaneous or recurrent motor branch of the median nerve, tendon adhesion, hypertrophic scarring and laceration of the superficial palmar arch. Other possible complications are post-operative infection, hematoma, arterial injury, stiffness, and reflex sympathetic dystrophy. Recurrence of the symptoms has been observed, in cases that the carpal tunnel release was unsuccessful, which can be a result of incorrect or incomplete diagnosis.

5.1.4 Medication

Non-steroidal anti-inflammatory drugs (NSAIDs), such as aspirin, ibuprofen and other pain relievers, are used to relieve the symptoms of CTS and to control the pain. Cortisone injections or corticosteroids in a pill form may also help reduce swelling in more severe cases of the syndrome [5].

The injections of anesthetic combined with a corticosteroid into or near the carpal tunnel can be used for diagnosis and therapy. There is the risk of injuring the median nerve with the needle by performing a direct injection into the carpal tunnel. Other risks include causing intratendinous injection and tendon rupture or dysesthesias. As in surgery, wrist splinting is also recommended after an injection. If the first injection is successful, a second injection can be performed after a few months, although if another injection is needed, surgery should be considered as the next option.

5.1.5 Ultrasound therapy

Ultrasound therapy is possibly beneficial for the management of CTS, although more studies are needed to confirm that. Some studies have shown that twenty sessions of carpal tunnel ultrasound therapy administered over approximately seven weeks contribute in significantly greater reduction of symptoms at two weeks, seven weeks, and six months [10]. On the other hand, other studies doubt the benefit of this particular treatment for CTS.

5.1.6 Physical Therapy

A physical therapist can prescribe exercises to strengthen the wrist and the hand. Alternative approaches that can help relieve CTS symptoms include massage, yoga, ultrasound, chiropractic manipulation, and acupuncture.

5.2 Traditional Physical Therapy

Exercises for mobility and grip strength and pain reduction are the first two steps for a patient with reported loss of functionality due of CTS. The most common exercises recommended for the

conservative management of the symptoms of CTS are tendon gliding of the finger flexor tendons and nerve gliding of the median nerve [11]. Studies that included measurements of the carpal tunnel pressure have shown that occasional exercise of active wrist and finger motion lasting about one minute lowers the pressure within the carpal tunnel [12]. Gliding exercises involve moving the fingers in a specified pattern to help the tendons and nerves glide more smoothly through the carpal tunnel. While there is some evidence that gliding exercises can help relieve symptoms when used alone, these exercises appear to work better in combination with other treatments such as splinting.

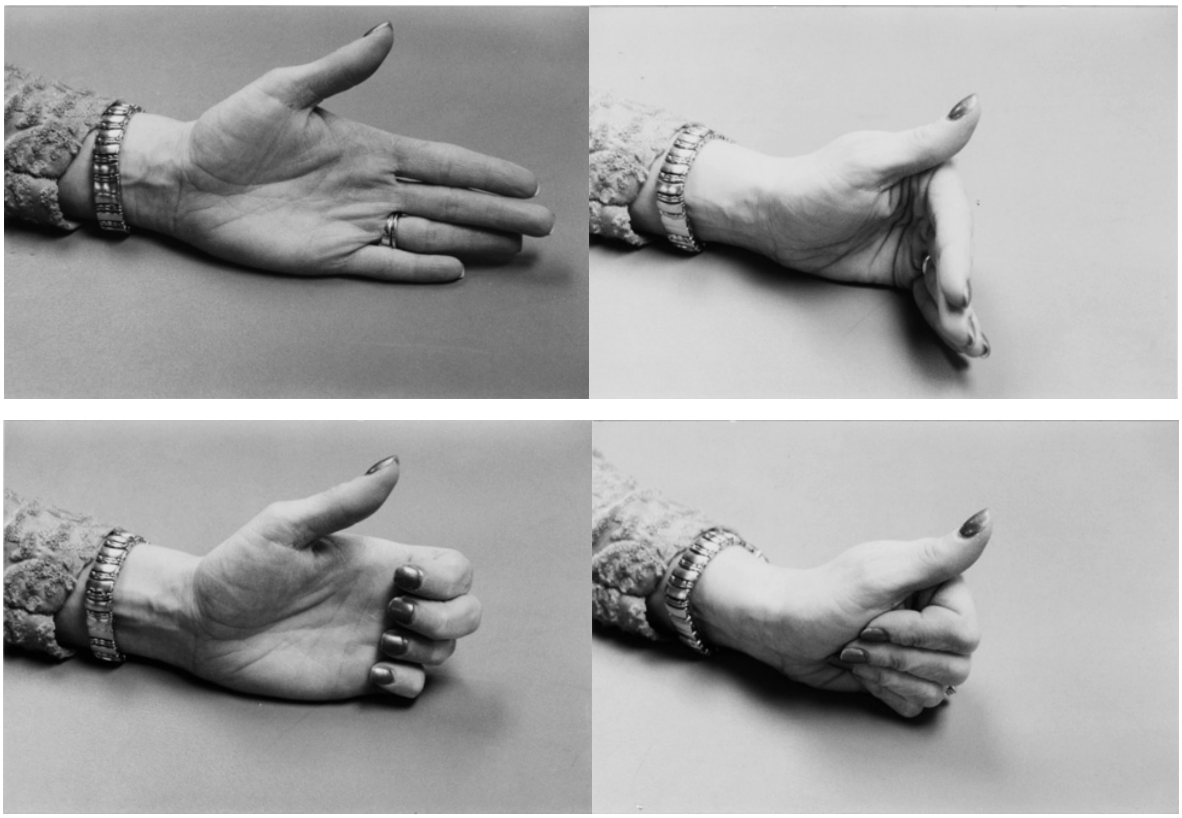


Figure 3: Common physiotherapy exercises for CTS treatment (adopted from [11])

Also manual therapy integrates a patented form of instrument-assisted soft tissue mobilization that gives the opportunity to the therapists to accurately detect and treat scar tissue and restrictions that effect normal function. Manual therapy techniques like active release technique and myofascial release are hands-on techniques that release tight tendons and musculature.

In the aggregate, physiotherapy is responsible for:

- Reducing pressure or inflammation, pain and loss of strength in your hand, wrist or arm
- Deciding whether a wrist splint is appropriate for the specific case
- Providing an individual program to improve functionality and increase the mobility of the wrist
- Instructing the patient to perform effective stretching exercises to prevent or minimize future symptoms appearance
- Providing specific advice on returning to work or normal everyday activities

5.3 Computer-aided, game-based therapy

A popular trend in today's game industry is to abandon conventional devices, such as keyboards, mice and joysticks in favor of mechanisms that detect the natural body motion, such as those employed by Nintendo's Wii and Microsoft's Kinect sensors. The purpose of these systems is to create the illusion to

the user of being a part of the virtual environment by translating natural body movements to game movements according to the game scenario. Interactive technologies like video games with motion-based input devices create an interesting alternative way to perform prescribed exercises at home via rich graphical and multimodal game contexts used to motivate players. As a result of exercising within meaningful game contexts, the quality of home-based self-initiated therapy is expected to improve.

Although computer games are considered a form of entertainment for younger people, their primary purpose can extend to much more than entertainment. The process of memorizing the rules of the game, of developing the tactics and making quick decisions helps players improve reflexes, cognition, attention, self-control, and motor abilities [13]. A special category of games, called serious games, have already been used as training simulation for professionals in the field of the military, health care and engineering. Games involving challenges and rewarding systems can motivate users to learn or perform different tasks for the purpose of the game that wouldn't be normally do on their own.

Therefore, serious games have also found application in the field of health care and help resolve or at least mitigate frequent health problems, such as severe mobility impairments, especially those that occur due to neurological and musculoskeletal problems. An example of such condition that causes deficiencies leading to mobility impairment and limitations in everyday activities is the multiple sclerosis, a chronic progressive neurological disease that affects the communication capacity of nerve cells in the brain and the spinal cord. Another example of such condition, which often creates similar difficulties in the movement and also usually contributes in the loss of balance, is the stroke. It involves disorders of the blood supply to the brain and subsequently leads to a partial loss of brain functionalities, since the affected area of the brain stops functioning. The aftermath of such condition can be decreased reaction time, the loss of range of motion, disordered movement organization and impaired force generation, thus affecting the patient's ability for self-sufficient and independent life [14].

In addition, an important benefit of the game-based physical therapy is that the technology allows the patients that require regular physical therapy sessions to perform prescribed exercises at home, so that they can avoid frequent visits to treatment centers. Furthermore, patients don't usually feel attached to the rehabilitation exercise program and lose their interest easily, especially when the results are not prompt, since the exercises consist of repetitive and monotonous tasks. Accordingly, it is very important to configure appropriate environments for the performance of the physical therapy exercises that will be able to motivate the patients [15] and facilitate the completion of repetitive, tedious tasks efficiently, so that the patients can stay dedicated to the program and have better recovery prospects [16]. In the case of CTS, the gliding exercises, which are also repetitive movements of the wrist and fingers that are commonly prescribed, can be successfully implemented in an engaging game environment with interesting features keeping the patient amused while performing therapy.

The last decade there have been several approaches [17], [18], and [19] to take advantage of the emerging motion sensor technology in the fields of physiotherapy and rehabilitation. The variety of the sensors and the constant optimization of their potentiality and features make them convenient for the therapy of different motor disabilities. Many of the attempts are mostly academic research [20] [21] that includes some prototypes of video games using sensors for physiotherapy treatment, but there are also commercial products developed by technological companies.

VirtualRehab Hands [14], a mini-gaming platform developed by Virtualware, is using the Leap Motion sensor to aid recovery from health problems such as Parkinson's disease and stroke (**Error! Reference source not found.**). According to Virtualware, their platform is among the first rehabilitation suites to

be classified as a medical device, under the EU's Medical Device Directives. The platform has been tested by patients and physiotherapists in installations in Europe, Latin America and the Middle East.

Figure 4: Virtual Rehab hand rehabilitation game (adopted from [14])

The games include exercises that help patients with motor disabilities and movement difficulties to preserve and regain upper and lower limb motor functionality. Additionally, the platform displays guidelines for the correct execution of the exercises and store the execution information or recording to provide feedback to the patient and allow the remote monitoring of the patients performance and progress by the therapist.

That same company has also developed the VirtualRehab-Body platform that employs the skeletal tracking and motion detection technology of Microsoft Kinect to detect and capture the movements of the patients within a game environment. The games include exercises that help patients with motor disabilities and movement difficulties to preserve and regain upper and lower limb motor functionality. Additionally, the platform displays guidelines for the correct execution of the exercises and store the execution information or recording to provide feedback to the patient and allow the remote monitoring of the patients performance and progress by the therapist.

The serious game "Kinect therapy – Boat Driving" developed by X-TECH Games team [22] utilizes the Kinect motion sensor to motivate the patients with different motor disabilities to perform their prescribed physical therapy exercises, but also urges casual users to train and gain body functionality improvement (**Error! Reference source not found.**). The game, involves the gathering of buoys in the sea by a boat driven by the user. The buoys score points and have to be gathered within certain time limits. These elements of time limit, score and rewarding instigate the users to execute the exercises often and for longer time. The user can operate the boat by performing rowing movements of the hands.

Figure 5: "Kinect therapy – Boat Driving" game by X-TECH Games team(adopted from [22])

Another commercial game platform that utilizes the Leap Motion controller for the purpose of rehabilitation and physical therapy is Visual Touch Therapy software program by the company Ten Ton Raygun [23]. The platform aims to make the therapy fun and engaging, while allowing the monitoring and assessment by the therapists to be easier. The target group of these therapeutic games is mostly patients that had stroke, but also individuals with spinal cord injuries, head injuries and nerve damage.

The Leap Motion Controller supports various movements and gestures, like swipes, taps, pushes and pulls, grabbing, grasping and side to side arm movements. Therefore, it gives the opportunity to the developers to implement a wide range of exercises. The games combine typical video game visual style, rewarding system with prompt feedback and various exercises for stroke rehabilitation, including hand and eye coordination challenges. The game scenario involves the Rocket Dog avatar that advances through the game levels and gains new abilities, such as speed, strength and rocket pack (**Error! Reference source not found.**).

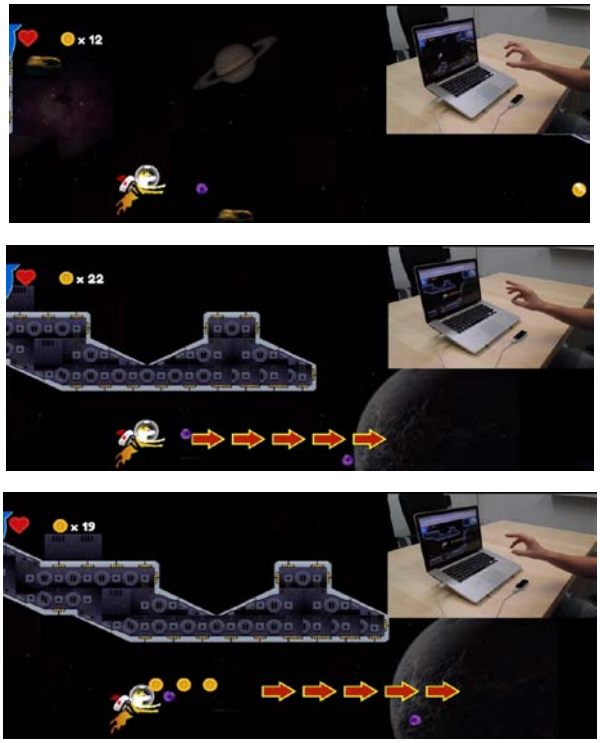


Figure 6: –"Rocket dog" stroke rehabilitation game that utilizes the Leap Motion sensor (adopted from [23])

As can be seen in the examples above, there are many attempts to develop physical therapy and rehabilitation platforms with the use of Kinect and Wii sensors. However, the Leap Motion sensor, despite the potential it offers in accuracy and portability, is still new in game development industry and especially in the field of therapeutic applications. The idea of developing a physical therapy game platform for the CTS with usage of Leap Motion controller seems very promising, due to increased accuracy of the sensor compared to other sensors, and quite original, since similar application examples haven't been found in the literature research.

REFERENCES

[1] "Physical Therapist's Guide to Carpal Tunnel Syndrome." [Online]. Available: http://www.moveforwardpt.com/symptomsconditionsdetail.aspx?cid=9f3cdf74-3f6f-40ca-b641-d559302a08fc. [Accessed: 05-Mar-2016].

[2] M. Chammas, J. Boretto, L. M. Burmann, R. M. Ramos, F. C. dos Santos Neto, and J. B. Silva, "Carpal tunnel syndrome – Part I (anatomy, physiology, etiology and diagnosis)," *Rev. Bras. Ortop. (English Ed.*, vol. 49, no. 5, pp. 429–436, 2014.

[3] "Carpal tunnel anatomy." [Online]. Available: https://www.mayoclinic.org/diseases-conditions/carpal-tunnel-syndrome/multimedia/carpal-tunnel-anatomy/img-20007899.

[4] "Carpal tunnel syndrome - Treatment." [Online]. Available: http://www.nhs.uk/Conditions/Carpal-tunnel-syndrome/Pages/Treatment.aspx. [Accessed: 01-Jan-2017].

[5] A. J. Viera, "Management of carpal tunnel syndrome," *Am. Fam. Physician*, vol. 68, no. 2, pp. 265–272, 2003.

[6] T. Gregory, D. Chamblin, J. Firestone, A. Friadman, H. Chirstopher, P. Douglas, M. Kliot, R. Lawrence, E. Thomas, B. Nicholas, D. Michael, T. Kjerulf, K. O'Bara, S. Carlson, and J. G, "Work-Related Carpal Tunnel Syndrome Diagnosis and Treatment Guideline," *Washingtonworkingsolutions.Net*. Washington State Department of Labor & Industries, pp. 1–18, 2017.

[7] M. W. Keith, V. Masear, K. C. Chung, K. Maupin, M. Andary, P. C. Amadio, W. C. WattersIII, M. J. Goldberg, R. H. HaralsonIII, and C. M. Turkelson, "Clinical Practice Guideline on the Diagnosis of Carpal Tunnel Syndrome," 2007.

[8] R. Garaka, "Nerve conduction study," *Health and Medicine*, 2012. [Online]. Available: https://www.slideshare.net/garakarabel/nerve-conduction-study.

[9] D. Kostopoulos, "Treatment of carpal tunnel syndrome: A review of the non-surgical approaches with emphasis in neural mobilization," *J. Bodyw. Mov. Ther.*, vol. 8, no. 1, pp. 2–8, 2004.

[10] A. H. Bakhtiary and A. Rashidy-Pour, "Ultrasound and laser therapy in the treatment of carpal tunnel syndrome," *Aust. J. Physiother.*, vol. 50, no. 3, pp. 147–151, 2004.

[11] S. L. Michlovitz, "Conservative Interventions for Carpal Tunnel Syndrome," *J. Orthop. Sport. Phys. Ther.*, vol. 34, no. 10, pp. 589–600, Oct. 2004.

[12] L. M. Rozmaryn, S. Dovelle, E. R. Rothman, K. Gorman, K. M. Olvey, and J. J. Bartko, "Nerve and tendon gliding exercises and the conservative management of carpal tunnel syndrome," *J. Hand Ther.*, vol. 11, no. 3, pp. 171–179, Jul. 1998.

[13] A. M. D. C. Souza, M. a Gadelha, E. a G. Coutinho, and S. R. Santos, "A video-tracking based serious game for motor rehabilitation of post-stroke hand impairment," *SBC J. 3D Interact. Syst.*, vol. 3, no. 2, pp. 37–46, 2012.

[14] "Virtualrehab." [Online]. Available: http://www.virtualrehab.info/.

[15] K. Lohse, N. Shirzad, A. Verster, N. Hodges, and H. F. M. Van der Loos, "Video Games and Rehabilitation," *J. Neurol. Phys. Ther.*, vol. 37, no. 4, pp. 166–175, Dec. 2013.

[16] F. Anderson, M. Annett, and W. F. Bischof, "Lean on Wii: Physical rehabilitation with virtual reality Wii peripherals," *Stud. Health Technol. Inform.*, vol. 154, pp. 229–234, 2010.

[17] A. Y. Wang, "Games for Physical Therapy," *SIGCHI Conf.*, pp. 4–8, 2012.

[18] M. E. Kho, A. Damluji, J. M. Zanni, and D. M. Needham, "Feasibility and observed safety of interactive video games for physical rehabilitation in the intensive care unit: a case series," *J. Crit. Care*, vol. 27, no. 2, p. 219.e1-219.e6, Apr. 2012.

[19] A. de Mauro, "Virtual Reality Based Rehabilitation and Game Technology," in *1st International Workshop on Engineering Interactive Computing Systems for Medicine and Health Care (EICSMed 2011), co-located*

with the ACM SIGCHI Symposium on Engineering Interactive Computing Systems (EICS 2011), 2011, pp. 48–52.

[20] M. Khademi, H. Mousavi Hondori, A. McKenzie, L. Dodakian, C. V. Lopes, and S. C. Cramer, "Free-hand interaction with leap motion controller for stroke rehabilitation," in *Proceedings of the extended abstracts of the 32nd annual ACM conference on Human factors in computing systems - CHI EA '14*, 2014, pp. 1663–1668.

[21] L. Geurts, V. Vanden Abeele, J. Husson, F. Windey, M. Van Overveldt, J.-H. Annema, and S. Desmet, "Digital games for physical therapy: fulfilling the need for calibration and adaptation," in *Proceedings of the fifth international conference on Tangible, embedded, and embodied interaction - TEI '11*, 2011, vol. 53, p. 117.

[22] "Kinect Physical Therapy – Boat Driving." [Online]. Available: http://x-tech.am/kinect-physical-therapy-boat-driving/.

[23] A. Gupta, "AXLR8R Spotlight: Making Physical Therapy Fun with Ten Ton Raygun," 2017. [Online]. Available: http://blog.leapmotion.com/axlr8r-spotlight-making-physical-therapy-fun-with-ten-ton-raygun/.

