Published in final edited form as:

Conf Proc IEEE Eng Med Biol Soc. 2013 ; 2013: 4625–4628. doi:10.1109/EMBC.2013.6610578.

Kinect-based Rehabilitation System for Patients with Traumatic Brain Injury

J. Venugopalan[2], **C. Cheng**[1], **T.H. Stokes**[2], and **May D. Wang**[1,2]

[1]Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA, 30332

[2]Wallace H. Coulter department of Biomedical Engineering, Georgia Institute of Technology, Atlanta, GA, 30332

Abstract

Traumatic brain injury is the leading cause of death and disability in the age group of 0 – 44 years. Though physical exercises have proven benefits in the rehabilitation process, the compliance rate of patients in the home environment is poor. In this paper we propose a system, MotionTalk, which captures and analyses motion data acquired using Microsoft Kinect. It gives a real time quantitative assessment of the exercises performed by TBI patients at home with respect to the same exercises performed in the clinic by utilizing relatively inexpensive contactless sensing and dynamic programming techniques. We compare this system to previous reminder systems, wearable systems, and motion capture systems. MotionTalk is less intrusive in nature and inexpensive to deploy at home because it is based on readily available hardware.

I. Introduction

Traumatic brain injury (TBI) is a major cause of death and disability worldwide, affecting nearly 2 million people annually with 30% in the United States of America [1]. In the United States of America, TBI accounts for nearly 30% of all injury related deaths. There are a total of 5.3 million patients in the US requiring rehabilitative assistance [1]. TBI is a disorder which can cause a host of physical, cognitive, social, emotional, and behavioral problems, and the outcomes can range from complete recovery to permanent disability or death. The medical field has made significant progress in the areas of diagnosis and mortality reduction [2, 3], but work still remains in the advancement of rehabilitation and quality of life for TBI patients. There is evidence that physical exercises are beneficial for cardiovascular fitness, mobility, and function following TBI, increased independence and the quality of life for the TBI patients [2]. However, the compliance rates for following instructions, including performing exercises, is an issue in TBI mainly due to problems of attention deficit [3].

There are relatively few studies in the area of the effective physical rehabilitation for TBI patients [4–6]. These studies mainly focus on the rehabilitation of individual areas such as upper body rehabilitation, gait and balance analysis. William et. al. performed gait analysis

Correspondence: May D Wang, Ph.D., Biomedical Engineering Department, Georgia Institute of Technology& Emory University, 313 Ferst Drive, UA Whittakker Building Suite 4106, Atlanta, GA, 30332, Tel: 404-3852954, Fax: 404-894-4243, maywang@bme.gatech.edu.

and found abnormalities in the gait of TBI patients with respect to the movement of the pelvis, knee and trunk, in addition to previously known abnormities in hip and ankle movements [7]. Ustinova et. al. conducted rehabilitation of the arm using a 3D game based approach, which was considered effective. Their system uses a standard camera and marker based motion capture system and the data analyses are performed using statistical methods, following dimensionality reduction [8]. These studies need expensive movement analysis equipment and trained personnel for setup and monitoring. Thus they are not suited for home based monitoring. The wearable sensing modalities for TBI are robotic devices for a specific set of exercises or accelerometer based sensing and classifying specific set of motions [9, 10]. We propose an inexpensive, easy to use system requiring minimal setup called MotionTalk, which is capable of full body physical rehabilitation using Microsoft Kinect. MotionTalk is a convenient and comfortable way to allow TBI patients the access to the expertise of physical therapists in the home environment. We analyze data captured from exercise sessions in comparison to data obtained from clinic sessions as template using dynamic time warping to give a quantitative score as a measure of the deviation.

II. System Design

The major components of the system are the Microsoft Kinect sensor, an adaptor for adding the Kinect as a peripheral to a PC, the data storage and analytics module, and the software interface.

A. Kinect Sensor

The Kinect sensor uses 2 video cameras and a NIR sensor to detect motion and distance. 20 joint locations and angles are accessible by using Microsoft Kinect SDK v1.5. We obtain XYZ coordinates & joint rotations as Quaternion (WXYZ) from 20 joints along with the time stamp at a maximal rate of 30 frames per second (fps).

B. Data Storage & Analytics Module

The data analytics module analyzes the information from Kinect obtained from home videos, compares it with the original video data from clinics and then arrives at a similarity score. The data from each session with the score and the template video pool of exercise templates are stored in a database.

C. Software Interface

The system has an interface with a panel to play the template video (clinic video) and a panel to show the user the exercise being performed in real time. The user can also see the joints tracked superimposed on either the video camera image or the infrared image. MotionTalk also provides feedback to the user and the therapists.

III. Data Analysis

Data analysis compares the captured test video (exercise performed by TBI patients at home) against the template video (exercise performed by TBI patients at the clinic) and gives the score as a measure of the deviation. The various steps involved in the analysis are illustrated in Figure 1.

A. Resampling

The data obtained from Microsoft Kinect is not uniformly sampled. This is due to the fact that data is made available along with the timestamp whenever tracking is complete. However, when template and test data are compared it is imperative that two are sampled at the same uniform sampling rate. Therefore the data obtained from Kinect is first resampled to 30fps.

Kinect data is multidimensional, i.e. (20 joints × (3) XYZ, (4) Quaternion, (1) timestamp [20 × 8]) per frame. (Here only the position value is used so it reduces to (20×4(XYZ, timestamp)) per frame. Resampling multidimensional data by interpolation traditionally uses multiple dimensional data to interpolate a single value (e.g. X, Y values to get intensity in images). However in this case we need to get all dimensions based on time alone. This was solved by interpolating each dimension separately against time using interpolation techniques. Spline interpolation techniques performed the best when tested against synthetic surfaces as compared to linear and cubic interpolations. The resampled data was then used to generate a feature space for further analysis.

B. Feature Extraction & Normalization

The features used were normalized position (20 × 3), velocity (20 × 3) and acceleration (20 × 3). The features obtained are in the coordinate frame in which the origin is at the center of the camera. This may lead to discrepancies when template and test are matched in scenarios where the subject locations are different in the two videos. To overcome this, normalization is done by taking all the positions with respect to the first frame. To account for the variations in distance, scaling is done by taking the distance between the two upper extremities in calibration pose (two hands stretched out) and the head and hip as the scaling factors for height and width.

C. Template generation & Updates

After feature extraction, template generation was performed to allow for comparison. Two types of templates were generated:

1) Static—These are templates which were generated from static poses such as the calibration pose. They are used as reference in normalization. Median value of all the frames in the static pose was used. The choice of median was arrived at after experimenting with mean, median and mode.

2) Moving—This is the template which serves as the reference video when the patient performs an exercise. The different approaches which were considered are combining 5 adjacent frames, using an aligned median of similar videos (similarity based on a threshold), weighted updates to the template based on distance.

D. Comparison with Template

Motion analysis studies have been used successfully in the study of diseases such as Parkinson's and Stroke [11–13] using techniques such as support vector machines, dynamic programming and Markov models. These studies were mainly involved in the classification of human motion. In our study we attempt to compare a test video (exercises performed at home) with a template video (exercises in the clinic) to give an assessment. Therefore we tried three different approaches to obtain scores. The different approaches tried are, direct frame by frame comparison, dynamic time warping and cross correlation to give the similarity score.

1) Direct Frame by Frame Comparison—This technique involves direct subtraction of the template from the test video and summing the deviations to give the score. This particular implementation was slightly modified to allow a simple alignment within 5 frames. If the deviations were below a threshold then they were accepted as a match. If not, they were rejected.

2) Dynamic Time Warping—Another technique used to compare the template and the test videos was dynamic time warping (DTW). The algorithm works by aligning the template video with the test based on Euclidean distance and then summing the deviation as the scores.

3) Cross Correlation—This technique gives the cross correlation score between test video and template video segments.

IV. Dynamic Time Warping Algorithm

Dynamic time warping works by aligning two signals with each other and finds the best path by minimizing the distance (Diff) between two signals. First, the two sequences A, B are placed on two axes of a matrix. Then a distance matrix D is calculated from the first matrix by minimizing equation 1.

$$\mathrm{D}(\mathrm{i},\mathrm{j}) = \min\left[\mathrm{D}(\mathrm{i}-1,\mathrm{j}-1), \mathrm{D}(\mathrm{i}-1,\mathrm{j}), \mathrm{D}(\mathrm{i},\mathrm{j}-1)\right] + \mathrm{d}(\mathrm{i},\mathrm{j}) \quad (1)$$

where d (i, j) is the cumulative distance at points i, j in the matrix. Dist equals D (M, N) for a matrix of dimensions M, N. There were three different variations which were tried out:

A. Method 1 - Spatial alignment

Each frame in the template was aligned to each frame in the test video using DTW and the sum was computed. This allowed spatial alignment in frame by frame basis. The score computed was the sum of individual scores. This was done to be able to align frames which were shifted spatially.

B. Method 2 - Smoothed spatial alignment

Frames were grouped using a moving window of length 5 and then DTW as performed to allow for spatial alignment. This was tested to smooth out minor variations between frames.

C. Method 3 - Temporal alignment

Video segments of 1 sec were aligned temporally to arrive at the DTW distances. The score was cumulative distances from all the video segments. This was done to be able to align frames which were shifted temporally.

D. Computation of Scores

The scores are arrved at by summing the distances obtained from individual comparisons (equation 2).

$$\mathrm{Score}=\sum_{k=0}^{N} D(k) \quad (2)$$

where N is the number of frames/segments of video depending on the method, Diff is the distance value obtained from each frame or segment. Therefore a smaller score means greater similarity and better exercise session as opposed to a greater score.

V. Results

A total of 16 videos captured with 6 different representative static and dynamic poses performed by 4 different volunteers (different number per person) in the age group 22–30 were analyzed. Figure 2(b) gives the same dynamic action (hand waving sideways) performed in red and different actions in blue. This was tested using direct comparison algorithm, DTW and cross correlation.

A. Cross-Correlation

Cross-correlation was tried as a fast and simple technique. However, the separation of the scores between similar and dissimilar videos was very small (Figure 3). The method could reliably distinguish similar and dissimilar videos only for static poses; hence all the diagonal values in figure 3 are not 1.

B. Direct Comparison

Figure 4 shows the scores which were obtained through the direct comparison algorithm implemented using each of the 16 videos as template and the other 15 as the test video. This algorithm picked up the motions performed at the same speed as the template as the same motion but failed to pick up the similar motion executed at different speeds. This method can be employed in situations where the speed of a particular action is critical.

C. Dynamic Time Warping

DTW was used as an improvement on the direct method to achieve a higher separation between the similar and dissimilar videos. A total of 3 variations were performed. Figure

5(a) gives the log of the DTW scores obtained using spatial alignment. This method was able to give a large score between most of the similar and dissimilar videos. This method failed to perform on a pair of videos as seen in the figure 5(a). Figure 5(b) gives the log of the DTW scores which were obtained using smoothed spatial alignment. This score separation of this method was better than method 1. Figure 5 (c) gives the log of the DTW scores which were obtained using temporal alignment. The score separation achieved by this method was similar to method 1. The scores produced from similar and dissimilar videos are large enough for the technique to be used in giving feedback to TBI patients and caregivers.

D. Test experiment design

The system will be tested on 10 patients recruited from Shepherd Centre, Atlanta, GA after IRB approval. TBI patients with mild injury requiring physical rehabilitation will be recruited in numbers of 2–3 per week. The system will be tested on 10 patients and the patients will be monitored through conventional methods of assessment. Once the results are obtained, the data will be analyzed for significant improvement in the physical condition as opposed to TBI patients using conventional methods of rehabilitation.

VI. Conclusion & Future work

The physical rehabilitation of TBI patients suffers from compliance related issues such as low motivation levels to exercise attributed to due to falls risk and attention deficit. MotionTalk is intended to improve patient engagement and assist physiotherapeutic rehabilitation in the home environment by analyzing 3-D body motions during physical therapy exercises and comparing them to 3-D motion videos, allowing therapists and patients to get a quantitative assessment. Then MotionTalk captures the patient’s movement data (20 joints, body segments & the angles between them) using Microsoft Kinect. The acquired motion is then analyzed using gesture recognition algorithms and the data analysis module assesses the exercise performance and provides scores. MotionTalk is proposed as a convenient and comfortable way to allow patients the access to the expertise of physical therapists in the home environment and to help reduce the health care costs through remote access and improved compliance. MotionTalk is currently limited in the evaluation of fine motor movements like hand and wrist movement. We will improve upon this in the future by adding more contactless sensors like the leap which will be out later this year.

In the future, we will evaluate the usefulness of the system on patients with TBI. In addition to a PC-based interface, MotionTalk will have a mobile-based app running on Android platform. We will also incorporate other features such as body orientation into motion analysis for improving the robustness. In the future, we will combine spatial and temporal alignment analysis. A weighted DTW approach will also be tried. Ultimately, we would like to see MotionTalk as a combination of open source motion sensing, cloud database, data analytics, and mobile app technologies, which provides a low-cost and easy-to-use channel for motion assessment.

Acknowledgments

The authors would like to thank Mr. Jason Lewitzke for aiding with the preliminary investigations in the project.

References

1. National Center for Injury Prevention and Control. Centers for Disease Control and Prevention; 2003. Available: www.cdc.gov/ncipc/wisqars
2. Schwandt M, Harris JE, Thomas S, Keightley M, Snaiderman A, Colantonio A. Feasibility and Effect of Aerobic Exercise for Lowering Depressive Symptoms Among Individuals With Traumatic Brain Injury: A Pilot Study. The Journal of Head Trauma Rehabilitationl. 2012; 27:99–103. DOI: 10.1097/HTR.0b013e31820e6858
3. Slifer, CLT Keith J.; Gerson, Arlene C.; Sevier, Robert C.; Kane, Alana C.; Amara, Adrianna; Clawson, Betsy P. Antecedent management and compliance training improve adolescents' participation in early brain injury rehabilitation. Brain Injury. 1997; 11:877–890. [PubMed: 9413622]
4. Cantin JF, McFadyen BJ, Doyon J, Swaine B, Dumas D, Vallee M. Can measures of cognitive function predict locomotor behaviour in complex environments following a traumatic brain injury? Brain Injl. Mar.2007 21:327–34.
5. Bland DC, Zampieri C, Damiano DL. Effectiveness of physical therapy for improving gait and balance in individuals with traumatic brain injury: A systematic review. Brain Injuryl. 2011; 25:664–679. 2011/07/01.
6. Cancelliere C, Cassidy JD, Cote P, Hincapie CA, Hartvigsen J, Carroll LJ, et al. Protocol for a systematic review of prognosis after mild traumatic brain injury: an update of the WHO Collaborating Centre Task Force findings. Syst Revl. 2012; 1:17.
7. Williams G, Morris ME, Schache A, McCrory PR. Incidence of Gait Abnormalities After Traumatic Brain Injury. Archives of Physical Medicine and Rehabilitationl. 2009; 90:587–593. 4//.
8. Ustinova KI, Leonard WA, Cassavaugh ND, Ingersoll CD. Development of a 3D immersive videogame to improve arm-postural coordination in patients with TBI. J Neuroeng Rehabill. 2011; 8:61.
9. Kexin X, Qi X, Jiping H, Yongji W, Zhongwei L, Xiaolin H. A wearable device for repetitive hand therapy. Biomedical Robotics and Biomechatronics, 2008. BioRob 2008. 2nd IEEE RAS & EMBS International Conference on. 2008:919–923.
10. Ince NF, Cheol-Hong M, Tewfik AH. In-Home Assistive System for Traumatic Brain Injury Patients. Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on. 2007:II-565–II-568.
11. Pastor I, Hayes HA, Bamberg SJM. A feasibility study of an upper limb rehabilitation system using kinect and computer games. Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE. 2012:1286–1289.
12. Sama A, Perez-Lopez C, Romagosa J, Rodriguez-Martin D, Catala A, Cabestany J, et al. Dyskinesia and motor state detection in Parkinson's Disease patients with a single movement sensor. Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE. 2012:1194–1197.
13. Chen Z, Wang L, Yung NHC. Adaptive human motion analysis and prediction. Pattern Recognitionl. 2011; 44:2902–2914.

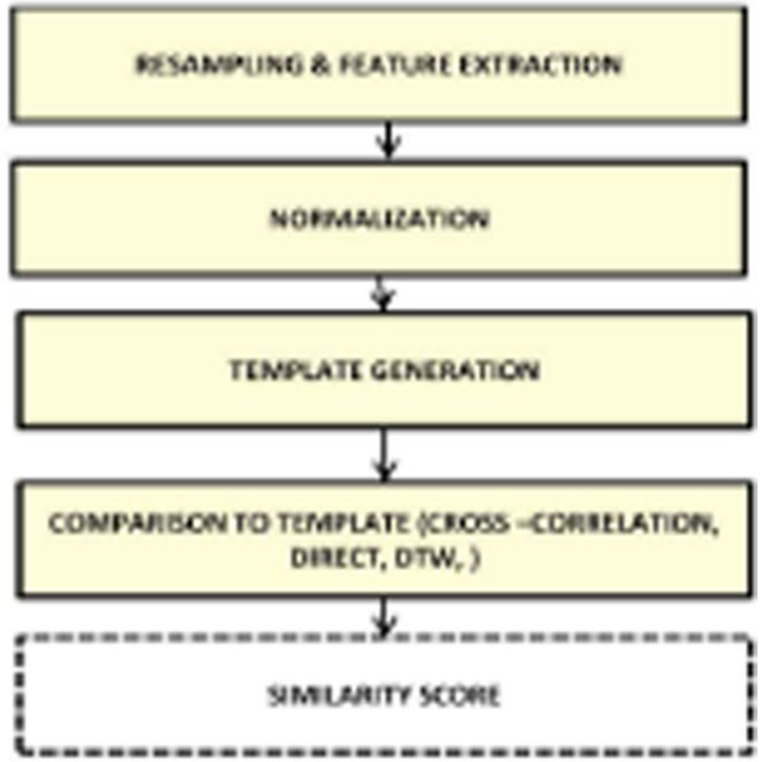

Figure 1.
Shins In Motion Analyst

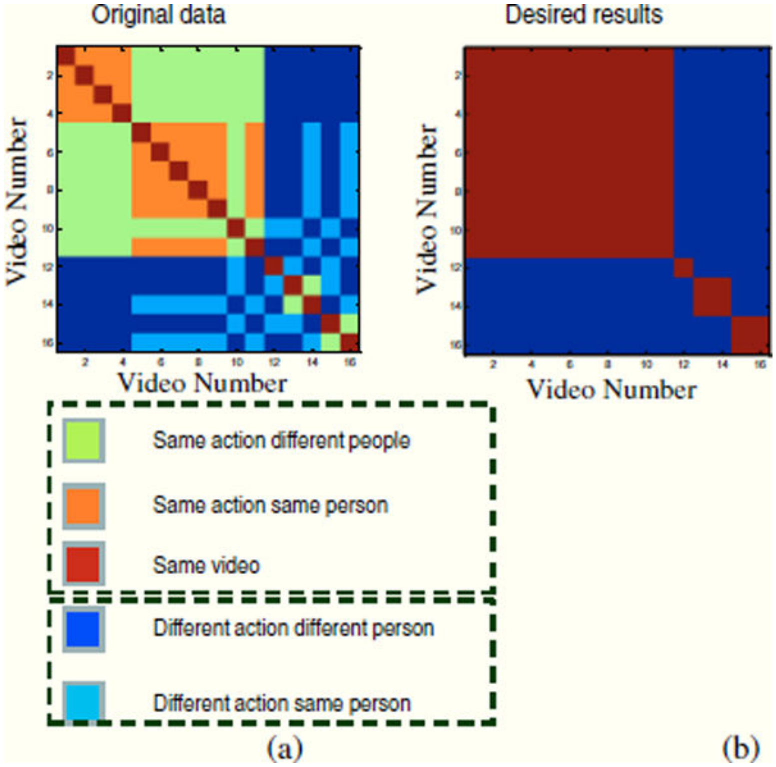

Figure 2.
(a) Actual data respresention (b) Desired results

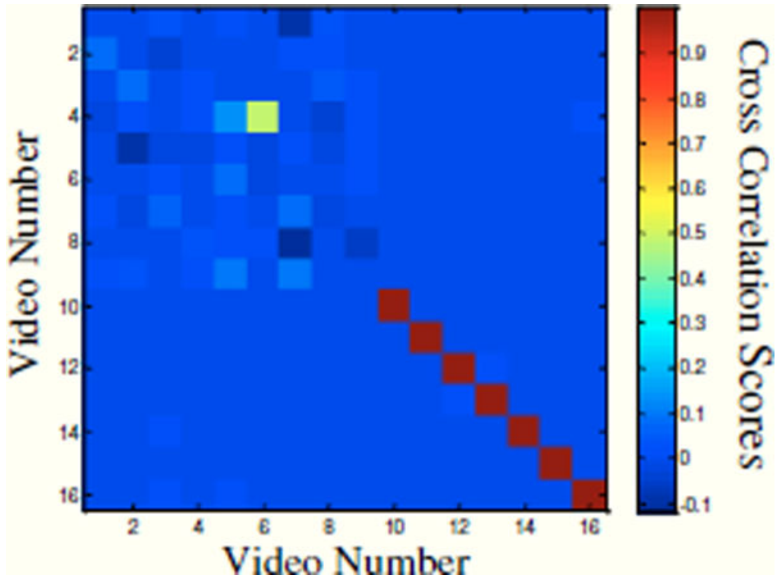

Figure 3.
Cross-correlation Scores by Comparing 16 Videos

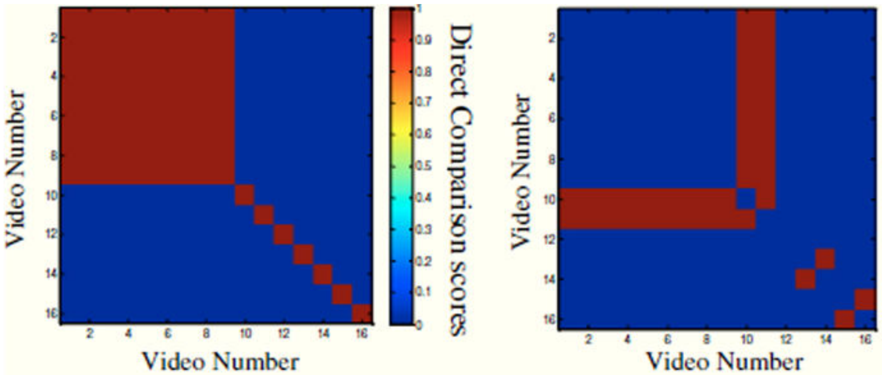

Figure 4.

Direct Comparison Results (a) Scores obtained from direct comparion for 16 videos (b) Difference from desired image

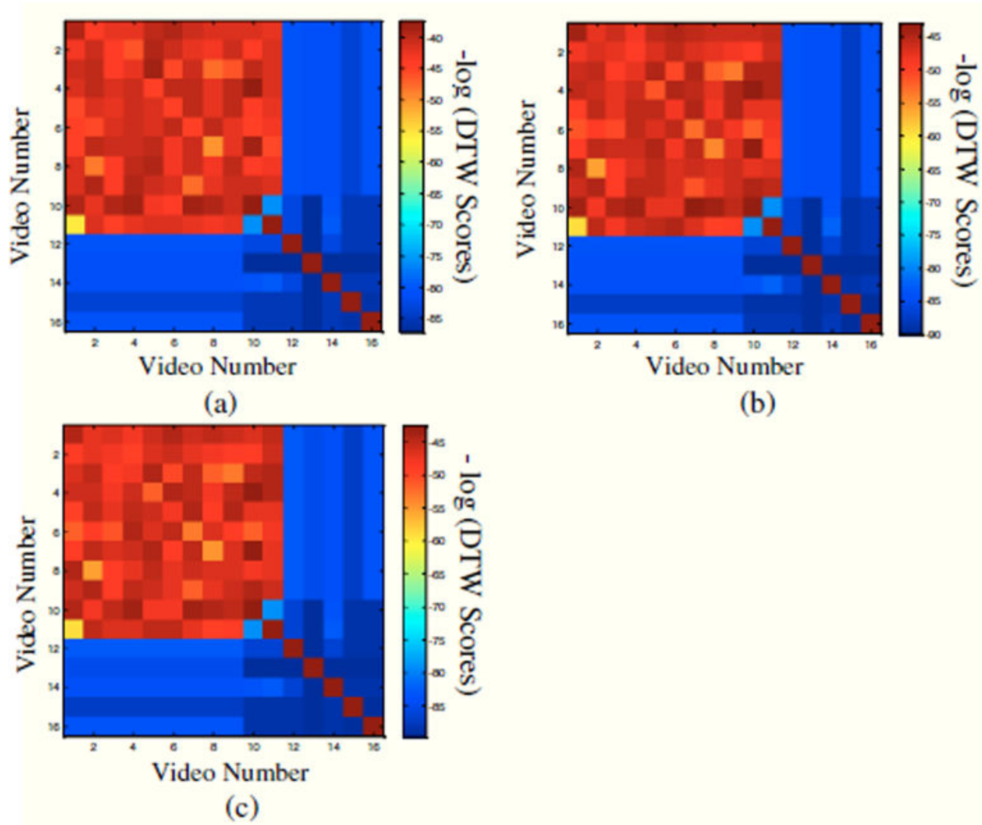

Figure 5.
Log of scores obtained from DTW for 16 videos (a) Method 1 – Spatial alignment (b) Method 2 – Smoothed Spatial alignment (c) Method 3 – Temporal Alignment