

Head Gesture Analysis using Matrix Group Displacement Algorithm

Mahmoud Z. Iskandarani

Department of Electrical Engineering, Faculty of Engineering,
Al-Zaytoonah University of Jordan, P.O. Box 911597, Post Code 11191, Amman, Jordan

Abstract: A novel algorithm for head gestures interpretation is designed and tested. The designed system carries out gesture detection and recognition using the MGDA algorithm, which implements random sampling and importance sampling, such technique can track head poses and estimate head positions. **Problem statement:** Head position is an important indicator of a person's focus of attention, which can be used as a key for multi-view face analysis assuming that face recognition and identification to be viewed dependently. This will help in selecting the best view model. Also, in the past few years face detection and person identification became important issues due to security concerns, leading to head gesture algorithm development and implementation. **Approach:** The captured image was allocated a map after which a file conversion process is carried out, allowing the next stage of image data conversion of head poses to be applied. This results in a specific number of matrices per pose that hold the necessary information. Such information was then allocated sequences representing head gesture poses which is combined for classification and correlation purposes to regenerate a predicted time reconstructed continuous movements. **Results:** A reliable, fast and robust approach for static head gesture recognition was achieved and presented. **Conclusion:** This very successful approach to head pose detection and gesture classification is strongly supported by its ability to correlate different signal input technologies as the devised algorithm can accommodate different inputs.

Key words: Computer vision, head gesture, motion tracking, head pose, MGDA system

INTRODUCTION

Head movement detection from an image sequences is a hard task due to variability in position, pose, size, orientation, color, expression, occlusion, resolution and lighting condition, hence, it is essential to develop efficient and adaptable algorithms to detect human faces. Detection of faces has been studied extensively over the past years, with various approaches. Gestures in general are expressive meaningful body motions with the intent to convey information or interact with the environment. Gesture recognition is the process by which gestures made by the user are acknowledged by an intelligence system (Villanueva and Cabeza, 2007; Srinivasa and Grossberg, 2008; Anishenko *et al.*, 2008; Murphy-Chutorian and Trivedi, 2009; 2010; Doshi, 2009; Junker *et al.*, 2008).

Understanding of human motion is especially important in fields such as biomechanics, robotics, human-computer interfaces, surveillance and computer animation. Based on the scope of each application area, various aspects of human motion have been studied from different perspectives. For example, motor control and learning theorists have studied how the central

nervous system creates and updates internal representations of limb dynamics, to perform complex, programmed movements in dynamic environments (Lin *et al.*, 2007; Eisenstein *et al.*, 2008; Suk *et al.*, 2010; Tonga *et al.*, 2007; Wu and Trivedi, 2008; Markin and Prakash, 2006; Choi and Kim, 2008; Bohme *et al.*, 2008). Sports analysts have studied the relationship between human physiology and movement efficiency. Ergonomics experts have modeled and employed complex relationships between human movement patterns, anthropometrics and ranges of motion to produce designs that better suit the human body. The fields of robotics and animation have developed algorithms and systems that mimic human motion and hence give us lifelike robots and humanoids. Methods for interpretation of communicative and expressive movements (such as facial movements, hand gestures and full body gestures) are needed for human-computer interaction paradigms and gesture-based recognition (Berenguer *et al.*, 2007; Littlewort *et al.*, 2006; Dornaika and Davoine, 2008; Bailenson *et al.*, 2008; Santhanam and Radhika, 2010; Toure and Beiji, 2010; Iskandarani, 2010; Khaddaj, 2010; Pavan *et al.*, 2010).

In this study a novel procedure for static head-pose estimation and a new algorithm for head gesture position are presented. Visual tracking is integrated into the novel MGDA system for measuring the position and orientation of a head. This system consists of interconnected modules that detect head position, provide initial estimates of the head's pose and can be implemented to continuously track head position and orientation.

MATERIALS AND METHODS

MGDA image-based gesture recognition system is used. The system comprises three main parts:

- Image processing and extraction of important clues, such as face or head position
- Tracking gesture features, such as face poses
- Gesture interpretation based on predefined gestures

Figure 1 illustrates the operation of MGDA.

Figure 2 shows images for practical (imperfect) head movements covering the five basic moves namely; left, right, straight, up, down.

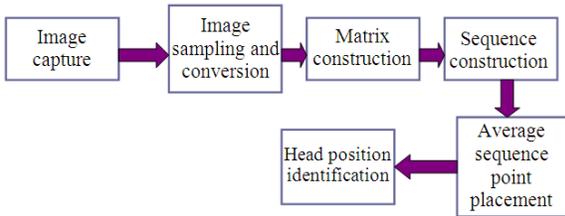


Fig. 1: MGDA system

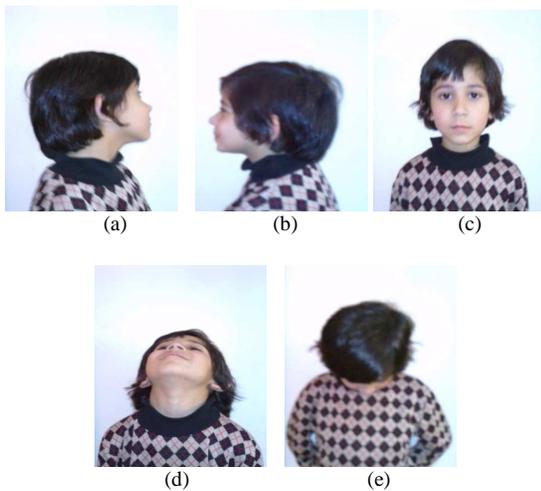


Fig. 2: Non-ideal head gestures performed by Eyad Iskandarani (2010)

RESULTS AND DISCUSSION

Figure 3-7 show discrimination made by the MGDA system after being applied to images from Fig. 2.

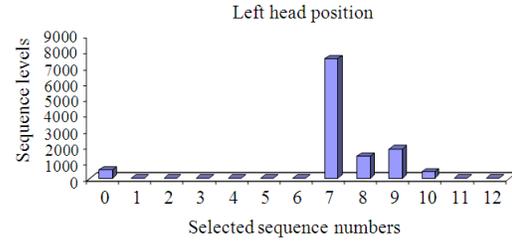


Fig. 3: Left head gesture

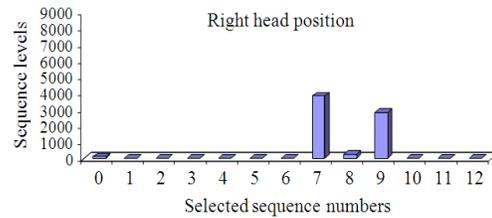


Fig. 4: Right head gesture

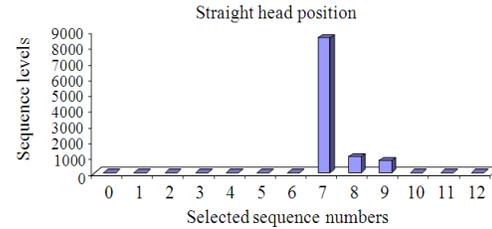


Fig. 5: Straight head gesture

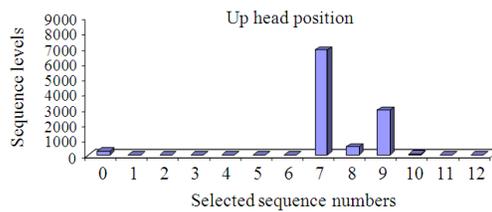


Fig. 6: Up head gesture

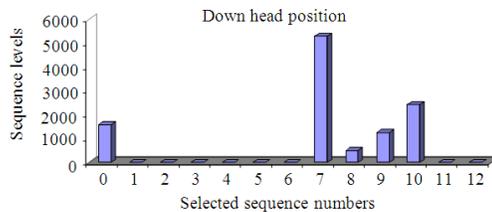


Fig. 7: Down head gesture

Figure 8 shows the obtained results for the five standard head movements. From point distribution, the following is realized:

- The algorithm successfully distinguished between different head gestures
- There is a clear relationship between the straight gesture and the other combinations [Left, Right], [Up, Down]
- For standard movements, the relationships for Sequence Levels via MGDA are given by:

$$\text{Straight} = \left(\frac{\text{Left} + \text{Right}}{2} \right) \quad (1)$$

$$\text{Straight} = \left(\frac{\text{Up} + \text{Down}}{2} \right) \quad (2)$$

Essentially, there are two basic movement sets:

- [Left→Right]
- [Up→Down]

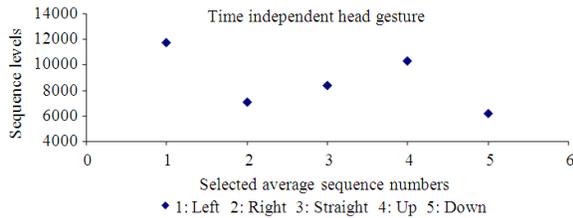


Fig. 8: Classified standard head gesture zones

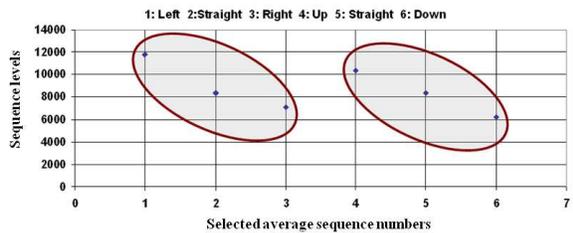


Fig. 9: Head gesture planes

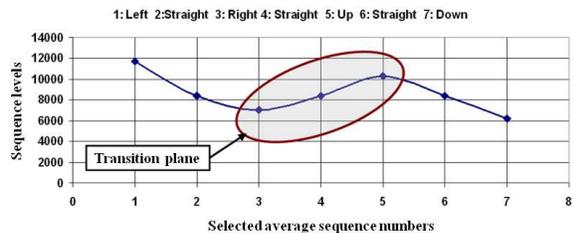


Fig. 10: Head gesture transitional plane

Even though the basic gesture sets are considered independent of each other; movements within each set must pass through the straight position. Figure 9 illustrate the two independent sets.

It is clear from Fig. 9 that there are two planes, one for [Up→Down] and the other for [Left→Right]. However, the movement translation should pass through centre (Straight) for standard movements or through a displacement of the centre point.

To move in between the two planes, a straight point position or its derivative is a requirement. The use of sequence numbers can be a substitute to discrete time intervals; hence, a transitional connection can be made between the two recognized planes as shown in Fig. 10.

The presence of two parallel planes describing Up→Down and Left→Right transitions leads to the following deductions:

- Between Planes:

$$\left(\frac{d\text{Straight}}{d\text{Right}} \right) = -f(\varphi) \left(\frac{d\text{Straight}}{d\text{Up}} \right)$$

- Inside Planes:

$$\left(\frac{d\text{Straight}}{d\text{Left}} \right) = -f(\theta) \left(\frac{d\text{Straight}}{d\text{Right}} \right)$$

$$\left(\frac{d\text{Straight}}{d\text{Up}} \right) = -f(\beta) \left(\frac{d\text{Straight}}{d\text{Down}} \right)$$

CONCLUSION

The presented Novel head pose scheme using MGDA that produces motion sequences is a feasible strategy for the identification of head poses and reconstruction of a sample continuous signal stream. It is demonstrated that such approach works well for static head images with non-ideal poses. The developed algorithm simplifies the process of acceptance rejection of non-relevant gestures.

REFERENCES

Anishenko, S., V. Osinov, D. Shaposhnikov, L. Podlachikova and R Comley *et al.*, 2008. Toward a robust system to monitor the head motions during PET based on facial landmarks detection: A new approach. Proceeding of the 2008 21st IEEE International Symposium on Computer-Based Medical Systems, June 17-19, IEEE Computer Society, Jyväskylä, pp: 50-52. DOI: 10.1109/CBMS.2008.19

- Bohme, M., M. Haker, T. Martinetz and E. Barth, 2008. A facial feature tracker for human-computer interaction based on 3D time of flight cameras. *Int. J. Int. Syst. Technol. Appli.*, 5: 264-273. DOI: 10.1504/IJISTA.2008.021289
- Bailenson, J., E. Pontikakis, I. Mauss, J. Gross and M. Jabon *et al.*, 2008. Real-time classification of evoked emotions using facial feature tracking and physiological responses. *Int. J. Hum.-Comput. Stud.*, 66: 303-317. DOI: 10.1016/j.ijhcs.2007.10.011
- Berenguer, E.P., C. Soria, O. Nasisi and V. Mut, 2007. Tridimensional pose estimation of a person head. *J. Phys.: Conf. Ser.*, 90: 1-8. DOI: 10.1088/1742-6596/90/1/012060
- Choi, S. and D. Kim, 2008. Robust head tracking using 3D ellipsoidal head model in particle filter. *Patt. Recog.*, 41: 2901-2915. DOI: 10.1016/j.patcog.2008.02.002
- Dornaika, F. and F. Davoine, 2008. Simultaneous facial action tracking and expression recognition in the presence of head motion. *Int. J. Comput. Vis.*, 76: 257-281. DOI: 10.1007/s11263-007-0059-7
- Doshi, A., 2009. A novel active heads-up display for driver assistance. *IEEE Trans. Syst. Man Cybernet. -Part B: Cybernet.*, 39: 85-93. DOI: 10.1109/TSMCB.2008.923527
- Eisenstein, J., R. Barzilay and R. Davis, 2008. Gesture salience as a hidden variable for co reference resolution and keyframe extraction. *J. Artif. Intell. Res.*, 31: 353-398. DOI: 10.1613/jair.2450
- Iskandarani, M.Z., 2010. Integrity classification algorithm of images obtained from impact damaged composite structures. *J. Comput. Sci.*, 6: 428-438. DOI: 10.3844/2010.428.438
- Junker, H., O. Amft, P. Lukowicz and G. Troster, 2008. Gesture spotting with body-worn inertial sensors to detect user activities. *Patt. Recog.*, 41: 2010-2024. DOI: 10.1016/j.patcog.2007.11.016
- Khaddaj, S., 2010. An integrated model for the versioning of image objects. *J. Comput. Sci.*, 6: 47-51. DOI: 10.3844/2010.47.51
- Lin, C.S., C.W. Hob, C.N. Chana, C.R. Chaua and Y.C. Wu *et al.*, 2007. An eye-tracking and head-control system using movement increment-coordinate method. *Opt. Laser Technol.*, 39: 1218-1225. DOI: 10.1016/j.optlastec.2006.08.002
- Littlewort, G., M.S. Bartlett, I. Fasel, J. Susskind and J. Movellan, 2006. Dynamics of facial expression extracted automatically from video. *Image Vis. Comput.*, 24: 615-625. DOI: 10.1016/j.imavis.2005.09.011
- Markin, E. and E.C. Prakash 2006. Tracking facial features with occlusions. *J. Zhejiang Univ.-Sci. A.*, 7: 1282-1288. DOI: 10.1631/jzus.2006.A1282
- Murphy-Chutorian, E. and M.M. Trivedi, 2009. Head pose estimation in computer vision: A survey. *IEEE Trans. Patt. Anal. Mach. Intell.*, 31: 607-626. DOI: 10.1109/TPAMI.2008.106
- Murphy-Chutorian, E. and M.M. Trivedi, 2010. Head pose estimation and augmented reality tracking: An integrated system and evaluation monitoring driver awareness. *IEEE Trans. Intell. Trans. Syst.*, 11: 300-311. DOI: 10.1109/TITS.2010.2044241
- Pavan, K.K., A.A. Rao, A.V.D. Rao and G.R. Sridhar Soung Y. 2010. Single pass seed selection algorithm for k-means. *J. Comput. Sci.*, 6: 60-66. DOI: 10.3844/2010.60.66
- Santhanam, T. and S. Radhika, 2010. A novel approach to classify noises in images using artificial neural network. *J. Comput. Sci.*, 6: 506-510. DOI: 10.3844/2010.506.510
- Srinivasa, N. and S. Grossberg, 2008. A head-neck-eye system that learns fault-tolerant saccades to 3-D targets using a self-organizing neural model. *Neural Networks*, 21: 1380-1391. DOI: 10.1016/j.neunet.2008.07.007
- Suk, H.I., B.K. Sin and S.W. Lee, 2010. Hand gesture recognition based on dynamic Bayesian network framework. *Patt. Recog.*, 43: 3059-3072. DOI: 10.1016/j.patcog.2010.03.016
- Tonga, Y., Y. Wangb, Z. Zhuc and Q. Jia, 2007. Robust facial feature tracking under varying face pose and facial expression. *Patt. Recog.*, 40: 3195-3208. DOI: 10.1016/j.patcog.2007.02.021
- Toure, M.L. and Z. Bei, 2010. Intelligent sensor for image control point of eigenfaces for face recognition. *J. Comput. Sci.*, 6: 484-491. DOI: 10.3844/2010.484.491
- Villanueva, A. and R. Cabeza, 2007. Models for gaze tracking systems *EURASIP J. Image Video Process.*, 2007: 1-16. DOI: 10.1155/2007/23570
- Wu, J. and M.M. Trivedi, 2008. A two-stage head pose estimation framework and evaluation. *Patt. Recog.*, 41: 1138-1158. DOI: 10.1016/j.patcog.2007.07.017