Vision-based Hand Gesture Recognition for Human-Computer Interaction

X. Zabulisy, H. Baltzakisy, A. Argyroszy

yInstitute of Computer Science
Foundation for Research and Technology - Hellas (FORTH)
Heraklion, Crete, Greece

zComputer Science Department, University of Crete
Heraklion, Crete, Grece

fzabulis,xmpalt,argyrosg@ics.forth.gr

1 Introduction

In recent years, research e®orts seeking to provide more natural, human-centered means of interacting with computers have gained growing interest. A particularly important direction is that of

the object of interest. The reason appears to be its local scope in combination

canvas [LWH02, WLH01]. Methods that relate to hand-driven control focus on the detection and tracking of some feature (e.g. the ¯ngertip, the centroid of the hand in the image etc) and can be handled with the information extracted through the tracking of these features. The second type of interaction involves the recognition of hand postures, or signs, and gestures. Naturally, the vocabulary of signs or gestures is largely application dependent. Typically, the larger the vocabulary is, the hardest the recognition task becomes. Two early systems

Clearly, because of the pixel-by-pixel image comparison, template matching is

network is tr

the image, based on optical °ow. The extremal points of the trajectory were detected and used as gesture classi¯cation features. In [CBA+96a], the 3D trajectories of hands are acquired by stereo vision and utilized for HMM-based learning and recognition of gestures. Di®erent feature vectors were evaluated as to their e±cacy in gesture recognition. The results indicated that choosing the

mand or message from the user to the system, require that the posture and motion of hands is recognized and interpreted. Early gesture recognition applications supported just a few gestures that signi¯ed some basic commands or concepts to the computer system. For example in [DP93], a monocular vision system that

control a slide presentation. The systems in [ZNG+04] and [CRMM00, MMR00], recognize a few hand postures for the control of in-car devices and non-safety systems, such as radio/CD, AC, telephone and navigation system, with hand

3 The Proposed Approach to Human-Robot Interaction based on Hand Gestures

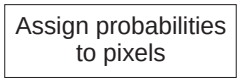
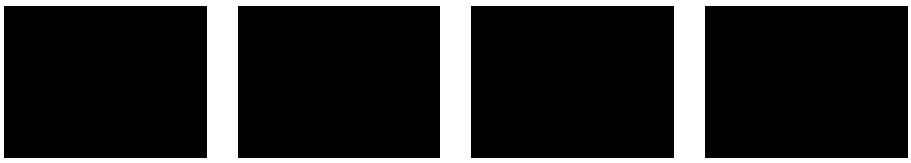


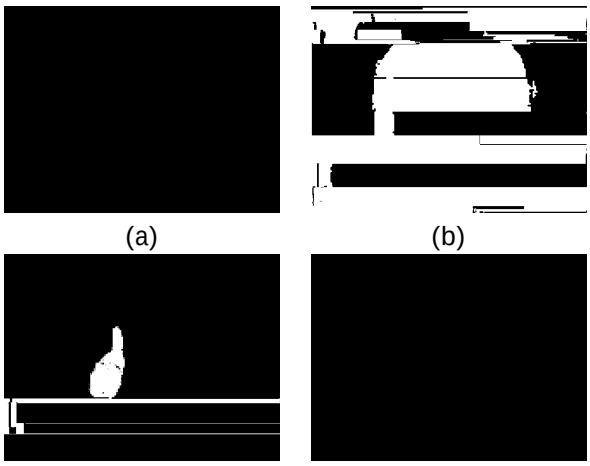
Fig. 1: Block diagram of the proposed approach for hand tracking and gesture



(a) (b) (c) (d)

Fig. 4:

3 The Proposed Approach to Human-Robot Interaction based on Hand G(ot)-342Ds



(a) (b)

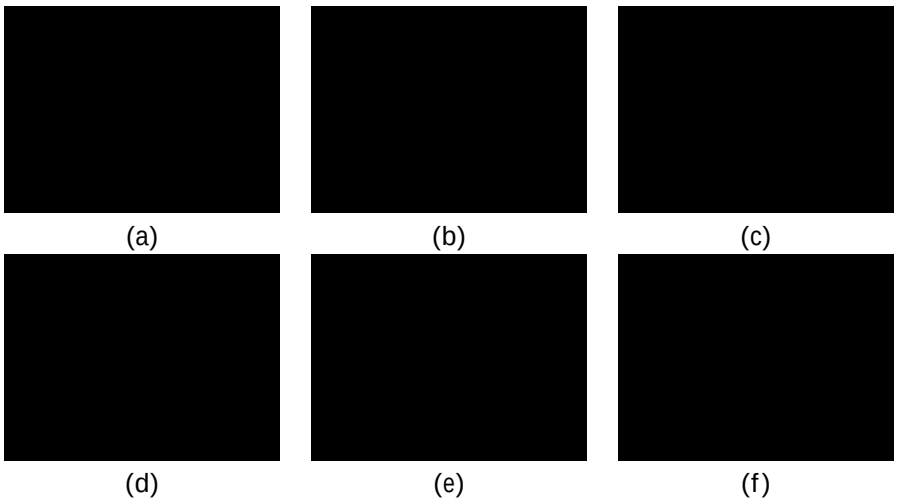


Fig. 7: Six frames of a sequence depicting a man performing gestures in an office environment.

Additionally, in this paper we have presented a novel gesture recognition system intended for natural interaction with autonomous robots that guide visitors in museums and exhibition centers. The proposed gesture recognition system

References

[ADS98] Y. Azoz, L. Devi, and R. Sharma. Reliable tracking of human arm dynamics by multiple cue integration and constraint fusion. In *Proc. IEEE Computer Vision and Pattern Recognition (CVPR)*, pages 905{910, Santa Barbara, CA, 1998.

[Ahm94] S. Ahmad. A usable real-time 3D hand tracker. In *Asilomar*

[BALT08] H. Baltzakis, A. Argyros, M. Lourakis, and P. Trahanias. Tracking of human hands and faces through probabilistic fusion of multiple visual cues. In *Proc. International Conference on Computer Vision SystemeALT08]*

[BHWS04] C. Bauckhage, M. Hanheide, S. Wrede, and G. Sagerer. A cognitive vision system for action recognition in o±ce environments. In *Proc. IEEE Computer Vision and Pattern Recognition (CVPR)*, volume 2, pages 827{833, 2004.

[BJ96] M. J. Black and A. D. Jepson. Eigentracking: Robust matching and tracking of articulated objects using a view-based representation. In *Proc. European Conference on Computer Vision*, pages 329{342, 1996.

[BJ98a] M. Black and A. Jepson. A probabilistic framework for matching temporal trajectories: Condensation-based recognition of gesture and expression. In *Proc. European Conference on Computer Vision*, volume 2, pages 909{924, 1998.

[BJ98b] M. Black and A. Jepson. Recognizing temporal trajectories us-

4 Summary

[CSW95] Y. Cui, D. Swets, and J. Weng. Learning-based hand sign recognition using shoslf-m. In *Int. Workshop on Automatic Face and Gesture Recognition*, pages 201{206, Zurich, 1995.

[CT98] R. Cutler and M. Turk. View-based interpretation of real-time optical °ow for gesture recognition. In *Proc. International Con-*

[DS92] Terzopoulos D. and R. Szeliski. *Tracking with Kalman Snakes*, pages 3{20. MIT Press, 1992.

[DS94a] J. Davis and M. Shah. Recognizing hand gestures. In *Proc. European Conference on Computer Vision*, pages 331{340, 1994.

[DS94b] J. Davis and M. Shah. Visual gesture recognition. *Vision, Im-*

[FR95]

[GWP99] S. Gong, M. Walter, and A. Psarrou. Recognition of temporal structures: Learning prior and propagating observation augmented densities via hidden markov states. In *Proc. International Conference on Computer Vision (ICCV)*, pages 157{162, 1999.

[HB96] G. Hager and P. Belhumeur. Real-time tracking of image regions with changes in geometry and illumination. In *Proc. IEEE Computer Vision and Pattern Recognition (CVPR)*, pages 403{410, Washington, DC, 1996.

[HCNP06]

[JP97] T. Jebara and A. Pentland. Parametrized structure from motion for 3D adaptive feedback tracking of faces. In *Proc. IEEE Com-*

[KL01] W. Kim and J. Lee. Visual tracking using snake for object's discrete motion. In *IEEE Int. Conf. on Robotics and Automation*, volume 3, pages 2608{2613, Seoul, Korea,, 2001.

[KM03] H. Kawashima and T. Matsuyama. Multi-viewpoint gesture [KL01]

[LL01] I. Laptev and T. Lindeberg. Tracking of multi-state hand models using particle ¯ltering and a hierarchy of multi-scale image features. In

[MI00] J. MacCormick and M. Isard. Partitioned sampling, articulated objects, and interface-quality hand tracking. In *Proc. European Conference on Computer Vision*, pages 3{19, 2000.

[ML97] P. Morguet and M. K. Lang. A universal HMM-based approach to image sequence classi¯cation. In *Proc. International Conference on Image Processing (ICIP)*, pages 146{149, 1997.

[MM95] D. J. Mapes and M. J. Moshell. A two-handed interface for object manipulation in virtual environments. *PRESENSE: Teleoperators and Virtual Environments*, 4(4):403{416, 1995.

[MM02] G. Mori and J. Malik. Estimating human body con¯gurations using shape context matching. In *Proc. European Conference on Computer Vision*, volume 3, pages 666{680, Copenhagen, Denmark, 2002.

[MMR00] G. McAllister, S. McKenna, and I. Ricketts. Towards a non-

[OZ97] R. O'Hagan and A. Zelinsky. Finger Track - a robust and real-time gesture interface. In *Australian Joint Conference on Artificial Intelligence*, pages 475–484, Perth, Australia, November 1997.

[Pet99] N. Peterfreund. Robust tracking of position and velocity with Kalman snakes. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 10(6):564–569, 1999.

[PHVG02] P. Pérez, C. Hue, J. Vermaak, and M. Gangnet. *Color-based probabilistic tracking. In European Conference on Computer Vision, pages 661–675, Copenhagen, Denmark, May 2002.*

[SH962]

[SHJ94b] J. Schlenzig, E. Hunter, and R. Jain. Vision based hand gesture interpretation using recursive estimation. In

[SSKM98] N. Shimada, Y. Shirai, Y. Kuno, and J. Miura. Hand gesture estimation and model re¯nement using monocular camera - ambiguity limitation by inequality constraints. In *IEEE Int. Conf. on Face and Gesture Recognition*, pages 268{273, Nara, Japan, 1998.

[ST05] L. Song and M. Takatsuka. Real-time 3D ¯nger pointing for an augmented desk. In *Australasian conference on User interface*, volume 40, pages 99{108, Newcastle, Australia, 2005.

[STTC06] B. Stenger, A. Thayanantl(re¯neou82(a)-346(S)823(or,)-350(And)-317(TR)-347(TCi)-287o
biasd Fhnd ralc27(akng)-31nitFer,7

[UO98]

4 Summary

[YAT02] M. H. Yang, N. Ahuja, and M. Tabb. Extraction of 2D motion trajectories and its application to hand gesture recognition. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 24(8):1061{1074, August 2002.

[Yin03] M. Yin, X.and Xie. Estimation of the fundamental matrix from uncalibrated stereo hand images for 3D hand gesture recognition. *Pattern Recognition*, 36(3):567{584, 2003.

[YK04] S. M. Yoon and H. Kim. Real-time multiple people detec-