Enhanced Hierarchical Model of Object Recognition Based on Saliency Map and Keypoint

Yan-Feng Lu*, Tae-Koo Kang**, Hua-Zhen Zhang*, Dong-Sung Pae* and ○Myo-Taeg Lim*

*School of Electrical Engineering, Korea University
**Dept. of Information and Communication Engineering, Sangmyung University

*(TEL: 02-3290-3804; FAX: 02-3290-3698; E-mail: {yanfeng, mlim}@korea.ac.kr)
** (TEL: 041-550-5405; E-mail: tkkang@smu.ac.kr)

Abstract Hierarchical Model and X (HMAX) presents an invariant feature representation, following the mechanisms of the visual cortex. Although HMAX in object recognition is robust, scale and shift invariant, it has been shown to be sensitive to rotational deformation. To address this, we propose a novel patch selection method saliency and keypoint based patch selection (SKPS). In addition, we suggest an SKPS based HMAX model (S-HMAX). In contrast to HMAX that employs the random patch deriving a significant amount of redundant information, S-HMAX uses SKPS to extract fewer numbers of features with better distinctiveness. To show the effectiveness of S-HMAX, we apply it to object categorization on TU Darmstadt (TUD) database. Experimental results demonstrate that the performance of S-HMAX is a significant improvement on that of conventional HMAX.

Keywords Object recognition, Classification, HMAX, Saliency map, Keypoint

1. Introduction

Object recognition has been widely used in the visual navigation of robots, and pedestrian detection [1]. In practice, the difficulties in recognition are basically caused by appearance variations of objects and background complexity in images. Object variability in terms of rotation, and illumination variations, especially in the case of cluttered backgrounds, seriously disrupts the recognition. To address this issue, Poggio *et al* presented an computational model of object recognition, called Hierarchical Model and X (HMAX) [2]. Although HMAX is robust to scale and shift deformations, it has been shown to be sensitive to rotational deformation [3], [4]. To improve its robustness to rotation, we propose a saliency and keypoint based HMAX model (S-HMAX).

The rest of the paper is organized as follows. In Section 2, we briefly review the HMAX model. In Section 3, we describe S-HMAX. In Section 4, we present experimental results based on TU Darmstadt (TUD) database. Finally, we give our conclusions.

2. HMAX review

HMAX is a computational framework with four layers: S1, C1, S2, and C2, as shown in Fig. 1. The framework follows the mechanisms of the primary visual cortex and builds feature representation by patch matching and maximum pooling operations.

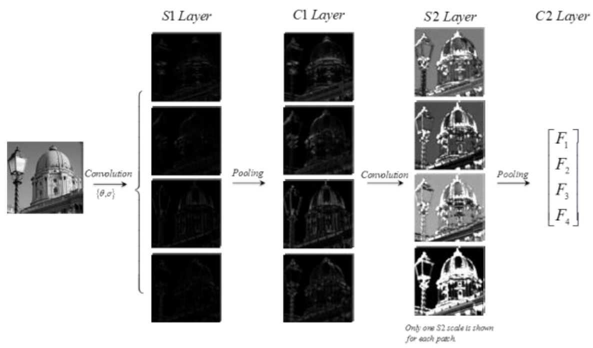


Fig.1. HMAX structure overview

3. S-HMAX model

3.1 SKPS

Because the stored patches in the C layers are the key components of the performance of HMAX, it is important to construct a proper patch set for classification task. Although a random selection of patches from the universal training set is an option, a refining process is preferable for visual systems. We propose the novel patch selection method SKPS that is based on saliency map and keypoints to extract fewer patches with better discrimination.

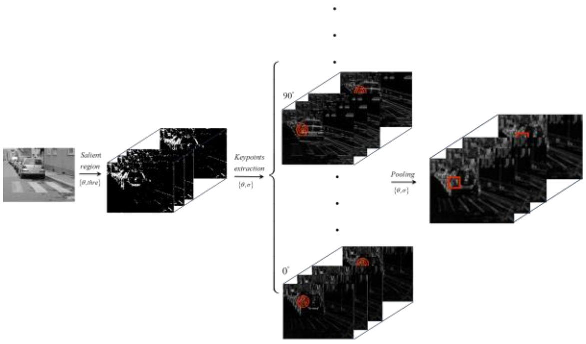


Fig. 2. Overview of SKPS in S-HMAX

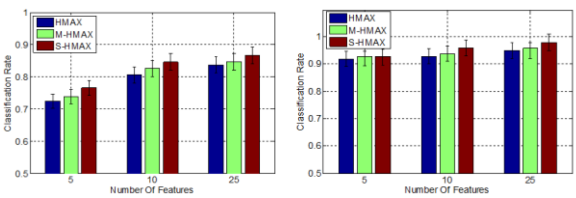
3.2 S-HMAX

In contrast to HMAX that selects patches randomly, in S-HMAX, we use SKPS to refine patches at C1 layers.

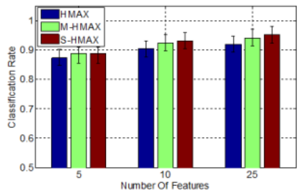
Because of the random selection, a significant amount of redundant information is extracted that degrades the performance. In S-HMAX, the patches are extracted from the salient region obtained by attention mechanism. The region contains more discriminative features. Within the region, we further find the locations with the most discriminative patches.

4. Experiment

The TUD is a public classification dataset. We used it to evaluate S-HMAX and draw comparisons with HMAX and a modified HMAX model named M-HMAX that uses a maximum energy based patch selection method [5].



(a) (b)



(c)

Fig. 3. Comparison of S-HMAX with standard HMAX and M-HMAX on the TUD database:
(a) Motorcycles, (b) cars, and (c) cows

5. Conclusion

In this article, we presented a saliency and keypoint based patch selection method aimed at reducing the redundant information in the HMAX model, and proposed S-HMAX with SKPS to extract discriminative and invariant features. Increasing the calculation speed of HMAX to real-time speeds will constitute our future work.

Acknowledgment

This work (NRF-2013R1A1A2008698) was supported by General Research Program through NRF grant funded by the Ministry of Education.

References

[1] Lee, D. Y., Lu, Y. F., et al.: ‘ 3D vision based local obstacle avoidance method for humanoid robot’, In Control, Automation and Systems (ICCAS), 2012 12th International Conference on , IEEE. pp. 473-475.
[2] Riesenhuber M, Poggio T.: ‘Hierarchical models of object recognition in cortex’, Nature Neuroscience, 1999, 2, (11), pp. 1019-1025.
[3] Serre T, Wolf L, Bileschi S, et al.: ‘Robust object recognition with cortex-like mechanisms’, Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2007, 29, (3), pp. 411-426.
[4] Lu Y F, Zhang H Z, Kang T K, et al.: ‘Extended biologically inspired model for object recognition based on oriented Gaussian–Hermite moment’. Neurocomputing, 2014, 139, pp. 189-201.
[5] Huang K, Tao D, Yuan Y, et al.: ‘Biologically inspired features for scene classification in video surveillance’, Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on, 2011, 41, (1), pp. 307-313.