Bulletin of Networking, Computing, Systems, and Software – www.bncss.org, ISSN 2186–5140
Volume 2, Number 1, pages 64–65, January 2013

Hand Pose Recognition Using Local Binary Patterns and Random Forests Classifier

Byungmin Kim, Sang-Heon Lee, Myoung-Kyu Sohn, Dong-Ju Kim, Hyunduk Kim
Dept. of IT-Convergence, Daegu Gyeongbuk Institute of Science & Technology (DGIST)
711-873, Daegu, Republic of Korea
{bmkim, pobbylee, smk, radioguy, hyunduk00}@dgist.ac.kr

***Abstract*—We propose an effective real-time hand pose recognition approach using local binary patterns and random forests classifier. Firstly, we localize the hand region from the entire image using the depth map from the depth camera. Then, we extract the feature vector from the hand image using local binary patterns. This feature vector is used to train a hand pose classifier based on random forests. In our experiments, we have constructed a large database of hand pose images and verified that the proposed LBP-based feature vector and random forests classifier is outperforms the other approaches**

Keywords—hand pose recognition; local binary pattern; random forests;

I. INTRODUCTION

Due to their convenience and naturalness, hand pose/gesture recognition methods are gaining attention as an upcoming complement of traditional input devices(keyboards, mice, joysticks, etc)[1]. However, gesture recognition remains a challenging problem. Digital gloves or marker-based methods provide feasible solutions, but wearing extra equipment on the hand(s) lessens the natural and general character of gesture interfaces[3].

Prior vision-based hand pose estimation techniques fall into two categories: appearance-based and model-based. Model-based methods match a hand model to features extracted from input images. Appearance-based methods are appealing for recognizing a small set of gestures. However, for larger gesture sets, such methods lack robustness and require extensive training data. "Estimation by Synthesis" is a combination of model-based and appearance-based methods. Graphical hand models are controlled to create images, which are then compared with input images to estimate hand poses. These methods often need to produce a large number of generated images to account for the variety of possible hand poses and views[2][3].

In Section 2-3, we review the local binary patterns and random forests briefly. Section 4 describes our experiments in detail. Finally, our conclusions and further works are discussed in Section 5.

II. LOCAL BINARY PATTERNS

The original LBP operator is a powerful means of texture description[4]. The operator labels the pixels of an image by thresholding the 3x3-neighbourhood of each pixel with the center value and considering the result as a binary number. Then the histogram of the labels can be used as a texture descriptor. Figure 1 shows an illustration of the basic LBP operator.

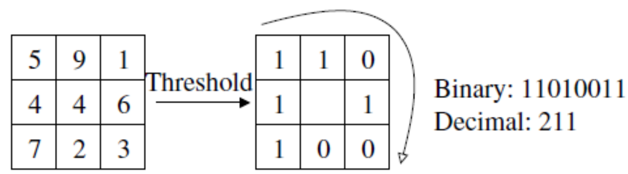

Figure 1. The basic LBP operator

Later the operator was extended to use neighborhoods of different sizes. Using circular neighborhoods and bilinearly interpolating the pixel values allow any radius and number of pixels in the neighborhood. For neighborhoods we will use the notation (P,R) which means P sampling points on a circle of radius of R. Another extension to the original operator uses so called uniform patterns. A Local Binary Pattern is called uniform if it contains at most two bitwise transitions from 0 to 1 or vice versa when the binary string is considered circular.

We use the following notation for the LBP operator: $LBP_{P,R}^{u2}$. The subscript represents using the operator in a (P,R) neighborhood. Superscript *u2* stands for using only uniform patterns and labeling all remaining patterns with a single label.

A histogram of the labeled image $f_l(x,y)$ can be defined as

$$H_i = \sum_{x,y} I\{f_l(x,y)=i\}, i=0,...,n-1, \tag{1}$$

in which n is the number of different labels produced by the LBP operator and

$$I\{A\} = \begin{cases} 1, \text{ A is true} \\ 0, \text{ A is false.} \end{cases} \tag{2}$$

This histogram contains information about the distribution of the local micro patterns, such as edges, spots and flat areas, over the whole image. The image is divided into regions $R_0, R_1, ...R_{m-1}$ and the spatially enhanced histogram is defined as

This work was support by the DGIST R&D Program of the Ministry of Education, Science and Technology of Korea (12-IT-03). It was also supported by Ministry of Culture, Sports and Tourism (MCST) and Korea Creative Content Agency (KOCCA) in the Culture Technology (CT) Research & Development Program(Immersive Game Contents CT Co-Research Center)

$$H_{i,j} = \sum_{x,y} I\{f_l(x,y) = i\} I\{(x,y) \in R_j\} \quad ,i = 0,...,n-1, j = 0,...,m-1 \tag{3}$$

The input hand region is segmented from the input image using depth map acquired from the depth camera. The hand region is divided into 4 sub-regions and the 4 histogram is constructed from each 4 sub-regions. The feature vector is made by concatenating the histograms.

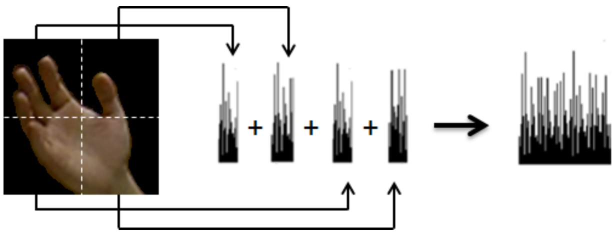

Figure 2. Histogram of LBP of hand image

III. RANDOM FORESTS

Random forest is an ensemble classifier that consists of many decision trees and outputs the class that is the mode of the classes output by individual trees. This is powerful tools capable of mapping complex input spaces into discrete or respectively continuous output spaces. A tree achieves highly non-linear mappings by splitting the original problem into smaller ones, solvable with simple predictors. Each node in the tree consists of a test, whose result directs a data sample towards the left or the right child. During training, the tests are chosen in order to group the training data in clusters where simple models achieve good predictions. Such models are stored at the leaves, computed from the annotated data which reached each leaf at train time. Breiman [5] shows that, while standard decision trees alone suffer from over-fitting, a collection of randomly trained trees has high generalization power. Random forests are thus ensembles of trees trained by introducing randomness either in the set of examples provided to each tree, in the set of tests available for optimization at each node, or in both[6].

We constructed the random forest using LBP-based feature vectors as the leaf node and 100 trees with depth 10.

IV. EXPERIMENTS

In order to assess the performance of our algorithm on realistic data, we constructed hand pose datasets. The datasets contains 175K images of 7 types of hand poses(one, two, three, close, open, left, right). We have used the depth and Color camera for image capturing. Figure 3 shows the samples of 7type of hand poses. The images have a resolution of 320x240 pixels, and a hand typically consists of around 150x200 pixels.

Figure 3. Sample images of hand poses

Firstly, we segmented the hand region from the entire image using the depth map. Then, we extracted the feature vector from the hand image using local binary patterns. This feature vector is used to train a hand pose classifier based on random forests.

Table 1 shows the success rate of the approaches for hand pose recognition. The performance was improved when the random forest classifier was employed.

TABLE I. EXPERIMENTAL RESULTS

Category	LBP+KNN	LBP+RF
2 category (open, close)	97.3	98.2
3 category (one, two, three)	94.4	95.9
4 category (open, close, left, right)	95.3	96.1
7 category (all)	93.8	94.3

V. CONCLUSIONS

In this work, we have presented an approach for hand pose recognition. In our experiments, the proposed approach achieved outstanding performance. In particular, the discriminative power is hardly decreased at all while the number of category is increased. Furthermore, because we do not rely on specific hardware like a GPU, our approach is also suitable for applications where hardware constraints limit the computational resources.

REFERENCES

[1] V. Pavlovic, R. Sharma, and T. Huang, Visual Interpretation of Hand Gestures for Human-computer Interaction: A Review, IEEE T. on Pattern Analysis and Machine Intelligence, vol. 19, pp. 677-695, 1997.

[2] A. Erol, G. Bebis, M. Nicolescu, R. D. Boyle, and X. Twombly, Vision-based hand pose estimation: A review, Computer Vision and Image Understanding, vol. 108, pp. 52-73, 2007.

[3] Z. Mo, Real-time Hand Pose Recognition Using Low-Resolution Depth Images, Computer Vision and Pattern Recognition, Vol. 2, pp. 1499-1505, 2006.

[4] T. Ojala, M. Pietikainen and T. Maenpaa, Multiresolution gray-scale and rotation invariant texture classification with local binary patterns, IEEE T. Pattern Analysis and Machine Inteligence, Vol. 24, No. 7, pp. 971-987, 2002.

[5] L. Breiman, Random Forests, Machine Learning, vol. 45, pp. 5-32, 2001.

[6] G. Fanelli, J. Gall and L. Van Gool, Real time head pose estimation with random regression forests, CVPR, pp. 617-624, 2011