Advances in Intelligent Systems and Computing

Volume 538

Series editor

Janusz Kacprzyk, Polish Academy of Sciences, Warsaw, Poland
e-mail: kacprzyk@ibspan.waw.pl

About this Series

The series "Advances in Intelligent Systems and Computing" contains publications on theory, applications, and design methods of Intelligent Systems and Intelligent Computing. Virtually all disciplines such as engineering, natural sciences, computer and information science, ICT, economics, business, e-commerce, environment, healthcare, life science are covered. The list of topics spans all the areas of modern intelligent systems and computing.

The publications within "Advances in Intelligent Systems and Computing" are primarily textbooks and proceedings of important conferences, symposia and congresses. They cover significant recent developments in the field, both of a foundational and applicable character. An important characteristic feature of the series is the short publication time and world-wide distribution. This permits a rapid and broad dissemination of research results.

More information about this series at http://www.springer.com/series/11156

Masato Akagi · Thanh-Thuy Nguyen
Duc-Thai Vu · Trung-Nghia Phung
Van-Nam Huynh
Editors

Advances in Information and Communication Technology

Proceedings of the International Conference, ICTA 2016

Springer

Editors
Masato Akagi
School of Information Science, Area of Human Life Design
Japan Advanced Institute of Science and Technology
Nomi-shi, Ishikawa
Japan

Thanh-Thuy Nguyen
Department of Computer Science
VNU University of Engineering and Technology
Hanoi
Vietnam

Duc-Thai Vu
Faculty of Information Technology
Thai Nguyen University of Information and Communication Technology
Thai Nguyen
Vietnam

Trung-Nghia Phung
Thai Nguyen University of Information and Communication Technology
Thai Nguyen
Vietnam

Van-Nam Huynh
School of Knowledge Science, Area of Knowledge Management
Japan Advanced Institute of Science and Technology
Nomi-shi, Ishikawa
Japan

ISSN 2194-5357 ISSN 2194-5365 (electronic)
Advances in Intelligent Systems and Computing
ISBN 978-3-319-49072-4 ISBN 978-3-319-49073-1 (eBook)
DOI 10.1007/978-3-319-49073-1

Library of Congress Control Number: 2016955568

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

This volume contains the papers presented at the International Conference on Advances in Information and Communication Technology (ICTA2016), which was held in Thai Nguyen city, Vietnam, during 12–13 December 2016. The conference was jointly organized by Thai Nguyen University of Information and Communication Technology (ICTU), Institute of Information Technology—Vietnam Academy of Science (IoIT), Fengchia University, Taiwan (FCU), Japan Advanced Institute of Science and Technology (JAIST) and National Chung Cheng University, Taiwan (CCU). The principal aim of ICTA2016 Conference is to bring together researchers, academics, practitioners and students in order to not only share research results and practical applications but also to foster collaboration in research and education in information and communication technology. The ICTA2016's Program Committee received a total of 150 submissions. Each submission was peer reviewed by at least two members of the Program Committee. Finally, 66 papers were chosen for presentation at Conference Sections and publication in the proceedings. Besides the main track, the conference has three invited speeches. We would like to express our appreciation to all the members of the Program Committee for their support and cooperation in this publication. We would like to thank Prof. Janusz Kacprzyk (Series Editor) and Dr. Thomas Ditzinger (Executive Editor, Interdisciplinary and Applied Science, Engineering, Springer) for their support and cooperation in this publication. We are also thankful to Anand Chozhan and his colleagues at Springer for providing a meticulous service for the timely production of this volume. Last but not the least, we wish to thank all the authors and participants for their contributions and fruitful discussions that made this conference a success.

December 2016 ICTA 2016

Masato Akagi
Thanh-Thuy Nguyen
Duc-Thai Vu
Program Chairs

Organization

Honorary Chairs

Bach Hung Khang	IoIT, Vietnam
Pham Viet Binh	ICTU, Vietnam

General Chairs

Nguyen Van Tao	ICTU, Vietnam
Nguyen Huu Cong	TNU, Vietnam

Program Chairs

Masato Akagi	JAIST, Japan
Nguyen Thanh-Thuy	VNU-UET, Vietnam
Vu Duc-Thai	LICTU, Vietnam

Publication Chairs

Huynh Van-Nam	JAIST, Japan
Phung Trung-Nghia	ICTU, Vietnam
Tran Dang Hung	Hanoi National University of Education, Vietnam
Nguyen Van Toi	ICTU, Vietnam

Organizing Chairs

Dang Thi Oanh	ICTU, Vietnam
Dam Thanh Phuong	ICTU, Vietnam
Le Hung Linh	ICTU, Vietnam
Ngo Manh Tuong	ICTU, Vietnam

Nguyen Hai Minh	ICTU, Vietnam
Nguyen Thanh Ha	TNU, Vietnam
Nguyen Van Nui	TNU, Vietnam
Nguyen Truong Thang	IoIT, Vietnam

Program Committee

Bui Thu Lam, Vietnam
Bui Ngoc Thang, Vietnam
Bui Quang Anh, Vietnam
Ching-Hwa Cheng, Taiwan
Chow Yen Desmond Sim, Taiwan
Chih-Hung Chou, Taiwan
Chyi-Ren Dow, Taiwan
Dang Quang A, Vietnam
Dang The Ngoc, Vietnam
Dang Thi Oanh, Vietnam
Dang Thanh Phu, Vietnam
Dang Van Duc, Vietnam
Dao Huy Du, Vietnam
Dao Phuong Nam, Vietnam
Dao Trung Kien, Vietnam
Do Nang Toan, Vietnam
Do Thanh Nghi, Vietnam
Do Thi Bac, Vietnam
Do Thi Ngoc Diep, Vietnam
Do Thi Tu Anh, Vietnam
Doan Nhat Quang, Vietnam
Doan Van Ban, Vietnam
Don-Gey Liu, Taiwan
Duong Chinh Cuong, Vietnam
Fang-Rong Hsu, Taiwan
Julia Tzu-Ya Weng, Taiwan
Gwo-Jiun Horng, Taiwan
Ha Duyen Trung, Vietnam
Ha Tien Ngoan, Vietnam
Hamido Fujita, Japan
Heng-Yi Su, Taiwan
Ho Dang Phuc, Vietnam
Ho Van Huong, Vietnam
Hoang Do Thanh Tung, Vietnam
Hoang Manh Thang, Vietnam
Huynh Van-Nam, Japan
Huynh Xuan Hiep, Vietnam
Kun-Chih Chen, Taiwan
Lai Hien-Phuong, Vietnam
Lai Khac Lai, Vietnam
Le Ba Dung, Vietnam
Le Duc Hau, Vietnam
Le Dung Muu, Vietnam
Le Hai Chau, Vietnam
Le Hoai Bac, Vietnam
Le Hoang Son, Vietnam
Le Hung Linh, Vietnam
Le Huy Tung, Vietnam
Le Quang Minh, Vietnam
Le Ngoc Tu, Japan
Le Nhat Thang, Vietnam
Le Thanh Ha, Vietnam
Le Thi Lan, Vietnam
Le Thi My Hanh, Vietnam
Le Trong Vinh, Vietnam
Le Trung Thanh, Vietnam
Le Viet Phuong, France
Luong Chi Mai, Vietnam
Luong The Dung, Vietnam
Masato Akagi, Japan
Nghiem Thi Phuong, Vietnam
Ngo Duc Thien, Vietnam
Ngo Hong Son, Vietnam
Ngo Quoc Tao, Vietnam
Ngo Thi Duyen, Vietnam
Ngo Xuan Bach, Vietnam
Nguyen Ai Viet, Vietnam
Nguyen Cat Ho, Vietnam
Nguyen Chien Trinh, Vietnam
Nguyen Cong Dieu, Vietnam
Nguyen Cuong, Vietnam
Nguyen Dai Hai, Vietnam

Nguyen Dinh Thuan, Vietnam
Nguyen Duc Dung, Vietnam
Nguyen Duc Nhan, Vietnam
Nguyen Duc Thuan, Vietnam
Nguyen Duy Minh, Vietnam
Nguyen Gia Nhu, Vietnam
Nguyen Hai Duong, Vietnam
Nguyen Hai Minh, Vietnam
Nguyen Hoai Nam, Vietnam
Nguyen Hoang Anh, Vietnam
Nguyen Hoang Nam, Vietnam
Nguyen Hoang Phuong, Vietnam
Nguyen Hong Quang, Vietnam
Nguyen Huu Cong, Vietnam
Nguyen Huu Quynh, Vietnam
Nguyen Kim Anh, Vietnam
Nguyen Linh Trung, Vietnam
Nguyen Long Giang, Vietnam
Nguyen Minh Tuan, Vietnam
Nguyen Nhi Gia Vinh, Vietnam
Nguyen Ngoc Linh, Vietnam
Nguyen Phuong Huy, Vietnam
Nguyen Quoc Cuong, Vietnam
Nguyen Quoc Khuong, Vietnam
Nguyen Quy Hy, Vietnam
Nguyen Son, Vietnam
Nguyen Tai Hung, Vietnam
Nguyen Thai Ha, Vietnam
Nguyen Thanh Hai, Vietnam
Nguyen Thanh-Thuy, Vietnam
Nguyen The Quang, Vietnam
Nguyen Thi Oanh, Vietnam
Nguyen Thi Phuong Giang, Vietnam
Nguyen Thi Thu Ha, Vietnam
Nguyen Thi Thuy, Vietnam
Nguyen Toan Thang, Vietnam
Nguyen Tran Hiep, Vietnam
Nguyen Truong Thang, Vietnam
Nguyen Van Chi, Vietnam
Nguyen Van Huan, Vietnam
Nguyen Van Huy, Vietnam
Nguyen Van Nui, Vietnam
Nguyen Van Tam, Vietnam
Nguyen Van Tao, Vietnam
Nguyen Van Thuy, Vietnam
Nguyen Van Toi, Vietnam
Nguyen Van Vy, Vietnam
Nguyen Viet Anh, Vietnam
Nguyen Viet Dung, Vietnam
Nguyen Viet Hung, Vietnam
Nguyen Xuan Huy, Vietnam
Nong Thi Hoa, Vietnam
Phan Thuong Cang, Vietnam
Pham Duc Long, Vietnam
Pham Minh Trien, Vietnam
Pham Ngoc Hung, Vietnam
Pham Quang Nhat Minh, Japan
Pham Thanh Giang, Vietnam
Pham Thanh Nam, Taiwan
Pham Viet Binh, Vietnam
Pham Xuan Nghia, Vietnam
Phan Xuan Hieu, Vietnam
Phung Trung-Nghia, Vietnam
Rakshtith Mysore Rajashekar, UK
Ren-Hung Hwang, Taiwan
Ta Tuyen, Vietnam
Tay-Jyi Lin, Taiwan
Teh-Lu Liao, Taiwan
Thai Quang Vinh, Vietnam
Tien-Yin Jimmy Chou, Taiwan
To Van Khanh, Vietnam
Tran Dang Hung, Vietnam
Tran Dinh Que, Vietnam
Tran Duc Tan, Vietnam
Tran Hoai Linh, Vietnam
Tran Hoang Tung, Vietnam
Tran Quy Nam, Vietnam
Tran Thanh Hai, Vietnam
Tran Thi Ngan, Vietnam
Tran Thu Ha, Vietnam
Tran Trung Duy, Vietnam
Tran Van Lang, Vietnam
Tran Xuan Nam, Vietnam
Truong Ha Hai, Vietnam
Truong Kien, USA
Truong Quoc Dinh, Vietnam
Vu Chan Hung, Vietnam
Vu Chien Thang, Vietnam
Vu Duc-Thai, Vietnam
Vu Duc Thi, Vietnam

Vu Duy Hai, Vietnam
Vu Duy Linh, Vietnam
Vu Duy Loi, Vietnam
Vu Hoai Linh, Vietnam
Vu Hoang Giang, Vietnam
Vu Manh Xuan, Vietnam
Vu Nhu Lan, Vietnam
Vu Tat Thang, Vietnam
Vu Tuyet Trinh, Vietnam
Vu Viet Vu, Vietnam
Vu Vinh Quang, Vietnam
Yao-Ting Huang, Taiwan
Yao-Tung Tsou, Taiwan
Yi-Chung Chen, Taiwan

Contents

An Evaluation of Hand Pyramid Structure for Hand Representation Based on Kernels

Van-Toi Nguyen[1(✉)], Thi-Lan Le[2], and Thanh-Hai Tran[2]

[1] School of Information and Communication Technology, Thai Nguyen University, Thai Nguyen, Vietnam
nvtoi@ictu.edu.vn

[2] International Research Institute MICA, HUST-CNRS/UMI-2954-GRENOBLE INP, Hanoi University of Science and Technology, Hanoi, Vietnam
{thi-lan.le,thanh-hai.tran}@mica.edu.vn
http://www.ictu.edu.vn, http://www.mica.edu.vn

Abstract. Hand posture recognition is an active research topic in computer vision and robotics with many applications ranging from automatic sign language recognition to human-system interaction. Recently, we have proposed a new descriptor for hand representation based on the kernel method (KDES) [1]. Our new descriptor inherits the main idea of KDES but we proposed three improvements to make it more robust. One of the improvements was that we introduced a new hand pyramid structure [14]. Intuitively, hand pyramid is more suitable to hand structure than conventional pyramid. In our previous work, we have demonstrated that the combination of improvements to KDES gives more accurate hand posture classification than using original KDES. However, it still lacks discussions and experimental evidences of the contribution of hand pyramid for hand representation. In this paper, we build specific hand dataset and conduct more experiments to show how hand pyramid contributes for hand representation. We will discuss deeply on the results and analyze the impact of this pyramid on hand posture classification.

Keywords: Hand posture recognition · Hand pyramid structure · Kernel Descriptor

1 Introduction

Vision-based hand posture recognition plays an important role in natural human-machine interaction. Hand posture recognition takes a hand region image as a result of hand detection step and returns a label of hand posture. Challenges to vision-based hand posture recognition are the following: (i) vision-based hand posture recognition is affected by changes in lighting condition, cluttered backgrounds, and changes in scale; (ii) hand is a deformable object; there exist a considerable number of mutually similar hand postures; (iii) applications using hand posture generally require real-time, user-independent recognition.

M. Akagi et al. (eds.), *Advances in Information and Communication Technology*,
Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_15

A number of hand recognition methods have been proposed to address these challenges [4,10,17,18]. These methods can be divided into two main categories depending how the hand is represented: explicit or implicit. Implicit representation relies on visual features which are computed directly from pixel values or reflect the relationship between pixels or regions. Simplest features are raw values of pixels. In this case, the image is often shaped into a 1D vector as the feature vector [12,13]. The using of raw pixel values makes high dimensionality of feature vector. Principal Component Analysis (PCA) is usually used [12,23] for dimension reduction. Instead of using directly pixel values, some other extract visual features that reflect relationship between pixels or regions. One of the popular features is extracted from Gabor filter [9]. Besides, many researchers also have employed Haar-like features [16,21,22]. Some other proposed to represent hands by orientation histogram [8], local binary pattern (LBP) [7] or SIFT [6] features.

The implicit hand representation methods do not reflect the structure of the hand. These methods hence do not take advantages of the specific structure of the hand. Some method using SIFT feature obtains good results on specific datasets that can provide a rich set of key points. However, when the resolution of hand image is low, the extracted set of key points is poor. In this situation, the method using SIFT feature will obtain a bad performance.

Explicit representation of hand bases on *intuitive* features. The methods belonging to this category often require good hand segmentation results to extract hand shape features (for example edges, contours) or topographical features such as fingers. Shape and topographical features are good for hand posture representation if we can segment well the hand region from the image [3,11,16,19,20]. However, hand segmentation is still a challenge in the real environment. Hand representation methods belonging to explicit presentation approach are intuitive and easy to understand for hand posture representation. However, these methods require a good hand segmentation result such as a clear hand contour. There was not a good combination between implicit and explicit representation. Some method used both of implicit and explicit features. Nevertheless, to use explicit features, we still need a good segmentation.

In our previous work [14], we proposed a method for hand presentation that is a flexible combination of implicit and explicit representation approaches. This hand representation method bases on kernel descriptor (KDES) [1]. We have made three improvements of KDES to make our descriptor more robust to orientation, scale change. One of the improvements was that we introduced a new pyramid structure which is intuitively more suitable to hand structure. However, we did not evaluate the impact of hand pyramid individually. In this paper, we build specific hand dataset and conduct more experiments to show how hand pyramid contributes for hand representation. We will discuss deeply on the results and analyze the impact of this pyramid on hand posture classification.

The remainder of the paper is organized as follows. The Sect. 2 reviews our method for hand posture recognition using the kernel method with hand pyramid structure. We then give some discussions and an evaluation of hand pyramid structure in the Sect. 3. The conclusions and directions for future work are given in the Sect. 4.

2 Hand Representation Using Hand Pyramid Structure

In [14], hand representation takes a hand region image (from now on called *image*, for short) as input and returns a descriptor of the hand candidate. It is composed of three sub-steps: Pixel-level feature extraction, Patch-level feature extraction, Image-level feature extraction.

2.1 Extraction of Pixel-Level Features

At pixel level, a normalized gradient vector is computed for each pixel of the image. The normalized gradient vector at a pixel z is defined by its magnitude $m(z)$ and normalized orientation $\omega(z) = \theta(z) - \overline{\theta}(P)$, where $\theta(z)$ is orientation of gradient vector at the pixel z, and $\overline{\theta}(P)$ is the dominant orientation of the patch P that is the vector sum of all the gradient vectors in the patch. This normalization will make patch-level features invariant to rotation. In practice, the normalized orientation of a gradient vector will be:

$$\widetilde{\omega}(z) = [sin(\omega(z))\ cos(\omega(z))] \tag{1}$$

2.2 Extraction of Patch-Level Features

Patch-level features are computed based on the idea of the kernel method. Derived from a match kernel representing the similarity of two patches, we can extract the feature vector for the patch using an approximate patch-level feature map, given a designed patch level match kernel function.

The gradient match kernel is constructed from three kernels that are gradient magnitude kernel $k_{\widetilde{m}}$, orientation kernel k_o and position kernel k_p. Gradient match kernel is defined as follows:

$$K_{gradient}(P, Q) = \sum_{z \in P} \sum_{z' \in Q} k_{\widetilde{m}}(z, z') k_o(\widetilde{\omega}(z), \widetilde{\omega}(z')) k_p(z, z') \tag{2}$$

where P and Q are patches of two different images that we need to measure the similarity. z and z' denote the 2D position of a pixel in the image patch P and Q.

The gradient magnitude kernel $k_{\widetilde{m}}$ is defined as:

$$k_{\widetilde{m}}(z, z') = \widetilde{m}(z)\widetilde{m}(z') \tag{3}$$

Where the normalized gradient magnitude $\widetilde{m}(z)$ is defined as:

$$\widetilde{m}(z) = \frac{m(z)}{\sqrt{\sum_{z \in P} m(z)^2 + \epsilon_g}} \tag{4}$$

where ϵ_g is a small constant. $m(z)$ is magnitude of the image gradient at a pixel z. The gradient magnitude kernel $k_{\widetilde{m}}$ is conspicuously a positive definite kernel.

Both the orientation kernel k_o and the position kernel k_p are Gaussian kernels which is of the form:

$$k(x, x') = \exp(-\gamma \|x - x'\|^2) \tag{5}$$

The factor γ will be defined individually for k_o and k_p that are denoted by γ_o and γ_p respectively.

Given the definition of match kernel, the approximate feature over image patch P is constructed as:

$$\overline{F}_{gradient}(P) = \sum_{z \in P} \widetilde{m}(z) \phi_o(\widetilde{\omega}(z)) \otimes \phi_p(z) \tag{6}$$

where $\otimes$ is the Kronecker product, $\phi_o(\widetilde{\omega}(z))$ and $\phi_p(z)$ are approximate feature maps for the kernel k_o and k_p, respectively. Given a match kernel function $k(x, y)$, the feature map $\varphi(.)$ for the kernel $k(x, y)$ is a function mapping a vector x into a feature space so as $k(x, y) = \varphi(x)^\top \varphi(y)$. Suppose that we have a set of basis vectors $B = \{\varphi(v_i)\}_{i=1}^{D}$, the approximation of feature map $\varphi(x)$ will be $\phi(x) = Gk_B(x)$, where G is defined by: $G^\top G = K_{BB}^{-1}$ and K_{BB} is $D \times D$ matrix with $\{K_{BB}\}_{ij} = k(v_i, v_j)$. k_B is a $D \times 1$ vector with $\{k_B\}_i = k(x, v_i)$.

The Kronecker product causes high dimension of the feature vector $\overline{F}_{gradient}$ (P). To reduce the dimension of $\overline{F}_{gradient}$, the kernel principal component analysis is applied into the joint basis vectors $\{\varphi_o(x_i) \otimes \varphi_p(y_j)\}_{i=1..d_o, j=1..d_p}$. Let t-*th* component ${\alpha^t}_{ij}$ is learned through kernel principal component analysis, following [1], the resulting gradient kernel descriptor for match kernel in (2) has the form:

$$\widetilde{F}^t_{gradient}(P) = \sum_{i=1}^{d_o} \sum_{j=1}^{d_p} \alpha^t_{ij} \sum_{z \in P} \widetilde{m}(z) k_o(\widetilde{\omega}(z), x_i) k_p(z, y_j) \tag{7}$$

2.3 Extraction of Image-Level Features Using Hand Pyramid Structure

Once patch-level features are computed for each patch, the remaining work is computing a feature vector representing the whole image. In [2], the authors used a spatial pyramid structure by dividing the image into cells using horizontal and vertical lines at several layers (Fig. 1(a)). This structure is generic, therefore does not take into account the specific shape of objects. In our work, as the hand is an object with a specific structure, we proposed a new pyramid structure specifically for the hand. In the following, we present in detail each step to build the final descriptor of the image.

Design a Hand Specific Pyramid Structure for Patch-Level Features Pooling: Fig. 1(b) shows the proposed hand pyramid structure. The main idea is to exploit characteristics of hand postures. Let the hand posture image have a size of $w \times h$. We observe that the regions at the image corners often do not contain information. For this reason, we only consider the area inside the inscribed

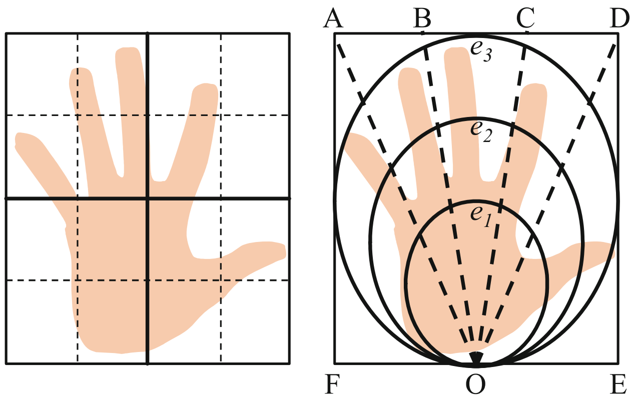

Fig. 1. (a) General spatial pyramid structure used in [2]. (b) The proposed hand pyramid structure.

ellipse of the hand image rectangle bounding box (e_3). The lines along the fingers converge at the lowest center point of the palm, near the wrist (O). Based on the structure of the hand, the ellipses (e_1, e_2, e_3) and the lines (OA, OB, OC, OD) are used to divide the hand region into parts that contain different components of the hand such as palm and fingers where $AB = BC = CD$. The detail of designed structure is described as: O is the midpoint of $FE(OF = OE)$. The ellipse e_1 is the inscribed ellipse of the rectangle that has a size of $(\frac{1}{2}w \times \frac{1}{2}h)$. The line FE is a tangent line of the ellipse e_1. The contact between the line FE and the ellipse e_1 is O. The ellipses are axis-aligned. In the similarity, the ellipsis e_2 is the inscribed ellipsis of the rectangle that has size of $(\frac{3}{4}w \times \frac{3}{4}h)$.

In a layer, we define a cell as being a full region limited by these ellipses and lines. In our work, the hand pyramid structure has 3 layers, (see Fig. 2).

- Layer 1: This layer contains only one cell defined by the biggest inscribed ellipse e_3.
- Layer 2: In [1], this layer has four rectangular cells. Unlike this, we create eight cells: three cells created from 3 ellipses and five cells created from the intersection of four lines with the biggest ellipse.
- Layer 3: This layer has 15 cells generated from the intersection between lines and three ellipses.

We can see that the hand pyramid gives a suitable representation for upright frontal hand postures in that the difference between postures is the configuration of the fingers (open or closed). In addition, the proposed hand representation has another advantage that is the image level feature is more invariant to slightly rotation/finger distortion. The reason of this ability is that the feature vector of a cell in pyramid is computed based on the set of feature vectors of patches belonging to the cell without concerning the locations of patches in the cell.

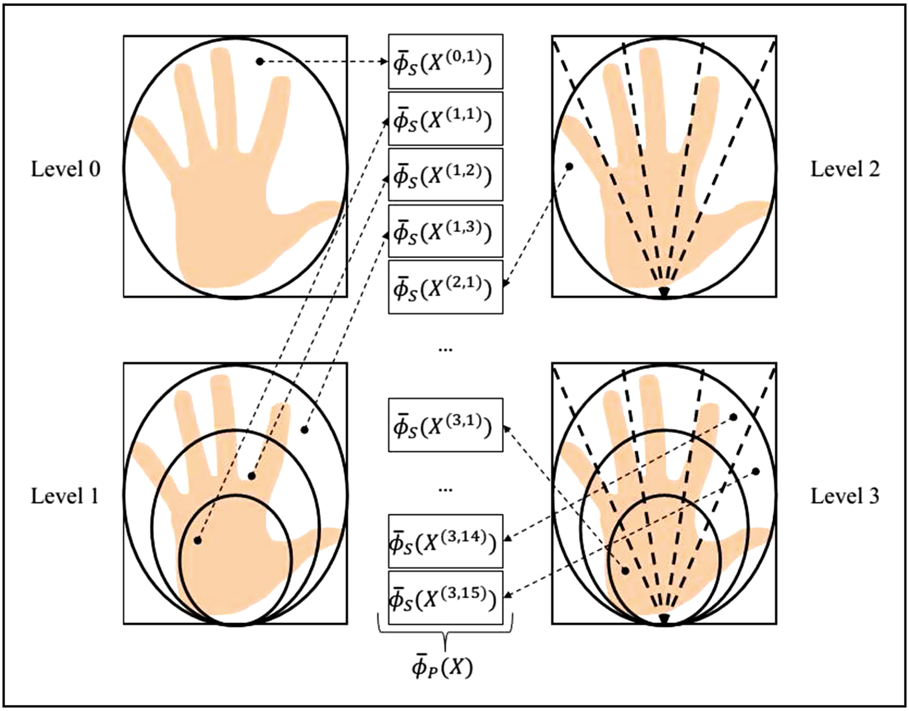

Fig. 2. Construction of image-level feature concatenating feature vectors of cells in layers of hand pyramid structure.

Moreover, the patch level feature is invariant to rotation. In case of our constraints (upright frontal hand postures with slightly rotation/finger distortion), the hand pyramid with normalized orientation gradient in the patch is suitable for hand representation, see Fig. 3(a,b). When hands strongly rotate or fingers heavily distort that make the same finger (or part of hand) in two hand images belongs to different cells in pyramid, above advantages of proposed hand presentation are not shown, see Fig. 3(c,d). In the case of strong rotation of hand, we could normalize hand image before applying hand pyramid.

Create the Final Descriptor of the Whole Image: To create the final descriptor of the whole image, we firstly use efficient match kernels (EMK) proposed in [1] to compute the feature vector for each cell of the hand pyramid structure, and then concatenate them into a final descriptor. Let C be a cell that has a set of patch-level features $X = \{x_1, ..., x_p\}$ then the feature map on this set of vectors is defined as:

$$\overline{\phi}_S(X) = \frac{1}{|X|} \sum_{x \in X} \phi(x) \tag{8}$$

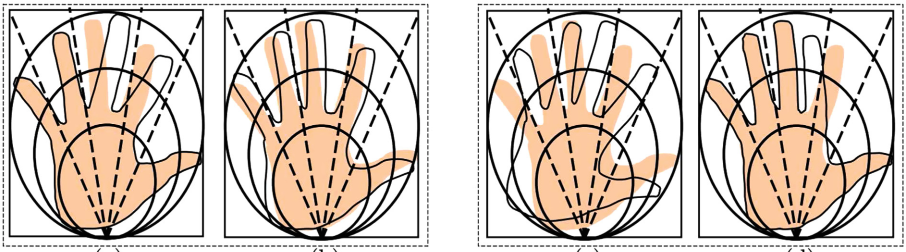

Fig. 3. (a) Slight finger distortion and (b) slightly hand rotation that do not make the same fingers belong to different cells; (c) heavy finger distortion and (d) strong hand rotation that make the same fingers belong to different cells.

where $\phi(x)$ is approximate feature maps for the kernel $k(x, y)$. The feature vector on the set of patches, $\overline{\phi}_S(X)$, is extracted explicitly.

Given an image, let L be the number of spatial layers to be considered. In this case $L = 3$. The number of cells in layer l-th is (n_l). $X(l, t)$ is set of patch-level features falling within the spatial cell (l, t) (cell t-th in the l-th level). A patch is fallen in a cell when its centroid belongs to the cell. The feature map on the pyramid structure is:

$$\overline{\phi}_P(X) = [w^{(1)}\overline{\phi}_S(X^{(1,1)}); ...; w^{(l)}\overline{\phi}_S(X^{(l,t)}); ...; w^{(L)}\overline{\phi}_S(X^{(L,n_L)})] \tag{9}$$

In (9), $w^{(l)} = \frac{\frac{1}{n_l}}{\sum_{l=1}^{L} \frac{1}{n_l}}$ is the weight associated with level l.

Figure 2 shows image-level feature extraction on the proposed hand pyramid structure. Until now, we obtain the final representation of the whole image, which we call image-level feature vector. This vector will be the input of a Multiclass SVM for training and testing.

3 Experimental Evaluation

3.1 Dataset

In order to evaluate the impact of hand pyramid structure on hand representation, we build a dataset that is a subset of L3i-MICA dataset [15]. This subset contains 13 upright frontal hand postures in that the difference between postures is the configuration of the fingers (open or closed), Fig. 4. In this dataset, hand posture images are captured in complex natural backgrounds. The hand regions are segmented manually. The number of training samples is 2859, and the number of testing samples is 2924. With this dataset, the accuracy we obtain in case of hand pyramid structure is 88.4 %, while the accuracy in case of general pyramid structure is 86.4 %. Hand pyramid structure makes the performance of hand posture recognition better (the accuracy increases 2 %). Thus we can remark that hand pyramid structure is suitable for upright frontal hand postures with open and closed fingers.

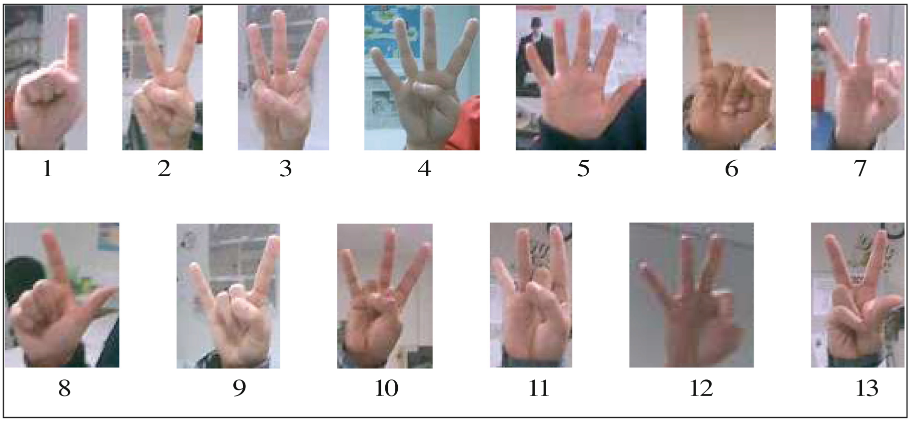

Fig. 4. Subset of L3i-MICA dataset.

4 Conclusions and Future Works

In this paper, we studied deeply the hand pyramid structure for hand representation in kernel method. We demonstrate empirically that our proposed hand pyramid is more suitable than general pyramid for hand posture recognition. This hand pyramid structure can be applied to detect open/closed fingers. In the future, we will evaluate hand pyramid structure with a more challenging dataset as well as a human-machine interaction using hand gestures.

Acknowledgement. This research is funded by Hanoi University of Science Technology under grant number T2016-PC-189.

References

1. Bo, L., et al.: Kernel descriptors for visual recognition. In: NIPS, pp. 1–9 (2010)
2. Bo, L., Sminchisescu, C.: Efficient match kernel between sets of features for visual recognition. In: NIPS, pp. 1–9 (2009)
3. Boughnim, N., Marot, J.: Hand posture recognition using jointly optical flow and dimensionality reduction. EURASIP J. Adv. Sig. Process. **2013**, 1–22 (2013)
4. Chaudhary, A., et al.: Intelligent approaches to interact with machines using hand gesture recognition in natural way: a survey. IJCSES **2**(1), 122–133 (2011)
5. Choi, J., Seo, B.: Robust hand detection for augmented reality interface. In: Proceedings of the 8th International Conference on Virtual Reality Continuum and its Applications in Industry, pp. 319–322 (2009)
6. Dardas, N.H., Georganas, N.D.: Real-time hand gesture detection and recognition using bag-of-features and support vector machine techniques. IEEE Trans. Instrum. Meas. **60**(11), 3592–3607 (2011)
7. Francke, H., Ruiz-del-Solar, J., Verschae, R.: Real-time hand gesture detection and recognition using boosted classifiers and active learning. In: Mery, D., Rueda, L. (eds.) PSIVT 2007. LNCS, vol. 4872, pp. 533–547. Springer, Heidelberg (2007). doi:10.1007/978-3-540-77129-6_47

8. Freeman, W.T., Roth, M.: Orientation histograms for hand gesture recognition. In: International Workshop on Automatic Face and Gesture Recognition, pp. 296–301 (1995)
9. Gupta, S., et al.: Static hand gesture recognition using local gabor filter. Procedia Eng. **41**, 827–832 (2012). Iris
10. Hasan, M.M., Mishra, P.K.: Hand gesture modeling and recognition using geometric features : a review. CJIPCV **3**, 1 (2012)
11. Kaufmann, B., Louchet, J., Lutton, E.: Hand posture recognition using real-time artificial evolution. In: Chio, C., Cagnoni, S., Cotta, C., Ebner, M., Ekárt, A., Esparcia-Alcazar, A.I., Goh, C.-K., Merelo, J.J., Neri, F., Preuß, M., Togelius, J., Yannakakis, G.N. (eds.) EvoApplications 2010. LNCS, vol. 6024, pp. 251–260. Springer, Heidelberg (2010). doi:10.1007/978-3-642-12239-2_26
12. Lockton, R., Fitzgibbon, A.W.: Real-time gesture recognition using deterministic boosting. In: BMVC, pp. 817–826 (2002)
13. Marcel, S., Bernier, O.: Hand posture recognition in a body-face centered space. In: Braffort, A., Gherbi, R., Gibet, S., Teil, D., Richardson, J. (eds.) GW 1999. LNCS (LNAI), vol. 1739, pp. 97–100. Springer, Heidelberg (1999). doi:10.1007/3-540-46616-9_9
14. Nguyen, V.-T., et al.: A new hand representation based on kernels for hand posture recognition. In: The 11th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2015) (2015)
15. Nguyen, V.-T.: Visual interpretation of hand postures for human-machine interaction (2015)
16. Ong, E.-J., Bowden, R.: A boosted classifier tree for hand shape detection. In: Sixth IEEE International Conference on Face and Gesture Recognition, pp. 889–894 (2004)
17. Pavlovic, V.I., et al.: Visual interpretation of hand gestures for human-computer interaction: a review. TPAMI **19**(7), 677–695 (1997)
18. Rautaray, S.S., Agrawal, A.: Vision based hand gesture recognition for human computer interaction: a survey. AIR **43**, 1–54 (2012)
19. Sgouropoulos, K., et al.: A dynamic gesture and posture recognition system. J. Intell. Robot. Syst. **76**(2), 283–296 (2013)
20. Stergiopoulou, E., Papamarkos, N.: Hand gesture recognition using a neural network shape fitting technique. EAAI **22**(8), 1141–1158 (2009)
21. Tran, T-T-H Nguyen, T-T-M : Invariant lighting hand posture classification. In: IEEE International Conference on Progress in Informatics and Computing (PIC), pp. 827–831 (2010)
22. Wachs, J., et al.: A real-time hand gesture system based on evolutionary search. In: GECCO2005: Tenth Genetic and Evolutionary Computation Conference, Washington, DC, USA (2005)
23. Wu, Y., Huang, T.S.: View-independent recognition of hand postures. In: Computer Vision and Pattern Recognition, pp. 88–94 (2000)