Available online at www.sciencedirect.com
ScienceDirect
Procedia Technology 10 (2013) 805 – 812
Procedia Technology

International Conference on Computational Intelligence: Modeling Techniques and Applications (CIMTA) 2013

Performance Analysis of Object Shape Classification and Matching from Tactile Images using Wavelet Energy Features

Shreyasi Datta[a], Anwesha Khasnobish[b,]*, Amit Konar[a], D.N. Tibarewala[b], R. Janarthanan[c]

[a]*Department of Electronics and Telecommunication Engineering,* [b]*School of Bioscience and Engineering,* [a,b]*Jadavpur University, India*
[c]*TJS College of Engineering, Chennai, India*

Abstract

Tactile images while grasping objects are acquired and wavelet based features are extracted for matching and classification. The performance of matching and classification is evaluated in terms of matching rate and classification accuracy along with the computation times. This comparison will help in determining the applicability of classification or matching in future works including real time applications. Highest classification accuracy is found to be 86%, in 0.0619sec, while the best matching rate obtained is 96% in 0.0041 sec. Thus Image matching is suitable for real time applications taking less computation time while providing significant performance improvement at the same time.

Selection and peer-review under responsibility of the University of Kalyani, Department of Computer Science & Engineering

Keywords: Tactile image; image matching; object-shape classification; wavelet energy; k-nearest neighbour; linear discriminant analysis; Naïve bayes; Euclidean distance.

1. Introduction

Touch sensation is essential for acquiring knowledge about the surroundings of an individual. It is important to incorporate artificial touch sensation in human computer interactive devices to enhance their controllability and

* Corresponding author.
E-mail address: anweshakhasno@gmail.com

Selection and peer-review under responsibility of the University of Kalyani, Department of Computer Science & Engineering
doi:10.1016/j.protcy.2013.12.425

produce human-like functionality. Tactile sensing is rapidly finding its application in various areas of robotics and rehabilitation. To identify and distinguish the objects around us, artificial tactile sensing system should be able to recognize object shape, size, texture, surface etc. There have been various methods for object shape recognition based on different techniques like the use of neural networks [1], [2], regional descriptors [3], image gradient [4], reconstruction of 3D object shape by fusion of visual and tactile information [5], fuzzy classification and reconstruction of 2-D shapes [6], use of Markov models for 2-D shape recognition [7] etc.

In the present work tactile images of 10 different object shapes are acquired from 30 different subjects and processed followed by feature extraction for matching as well as classification. Wavelet based features have been extracted from the acquired images. Image matching is an important image analysis and recognition technique. Image matching has been implemented through the use of Euclidean distance as a similarity measure. Classification is performed in one-against-one (OVO) basis. Three different classifiers were used, namely k-nearest neighbor (kNN), Naïve Bayes classifier, and Linear Discriminant Analysis (LDA). The performance of matching and classification is evaluated in terms of matching rate and classification accuracy respectively along with the computation times for object shape recognition from tactile images. This comparison will help in determining the applicability of classification or matching in future works including real time applications without compromising with the efficiency of the whole system. Object shape recognition is performed using single grasping instead of dynamic exploration. This reduces the time of tactile image acquisition as well as the computational complexity of processing a very large number of tactile images resulting from dynamic exploration, thereby making it suitable for real time applications of object shape identification using tactile sensors. The overall approach is presented in Figure 1.

Experimental paradigm is explained in section 2, followed by the methodology in section 3. Experimental results are given in section 4. The concluding remarks are outlined in section 5.

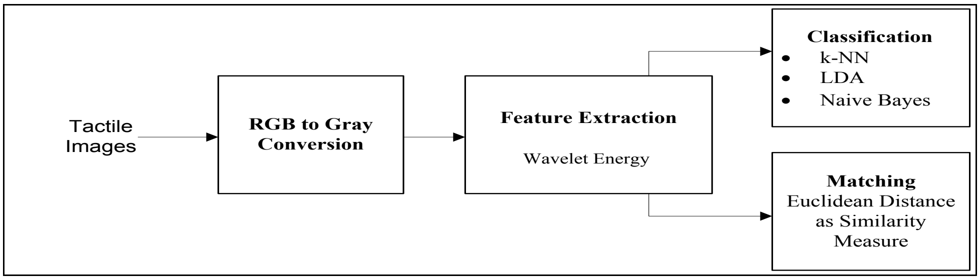

Figure 1. Flowchart depicting the course of work

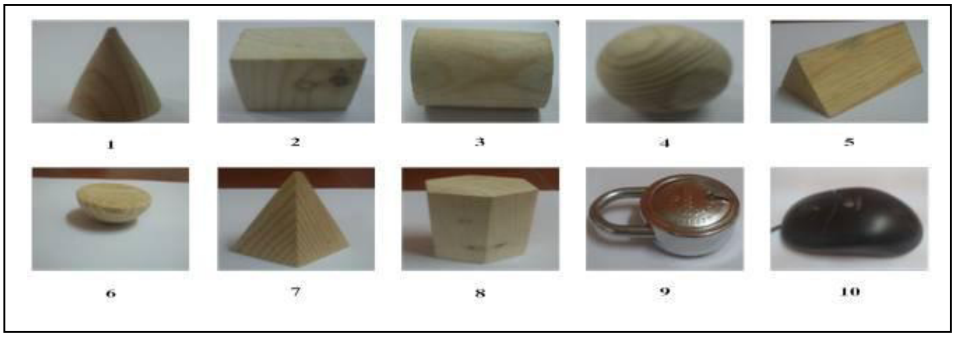

Figure 2. Objects used in the experiments

2. Experimental Paradigm

2.1. Tactile Sensor

The tactile images used in our experiment have been acquired by a capacitive MEMs based pressure sensor: the PPS TactArray [8] which has of 1024 tactile sensing elements arranged in a 32×32 grid. The MEMs based approach can provide better spatial resolution but cannot tolerate large forces as they are inherently fragile in nature [9]. In this work we are working with grip forces only while grasping different objects and the exerted pressures are very less and the flexible MEMs type sensor provides effective means to tactile image acquisition.

2.2. Objects

We have considered 10 different types rigid objects in our experiment. Out of these 10 objects, 8 are of generic shapes and the last two are non-generic household objects as shown in the Figure 2.

2.3. Subjects

Experiments were conducted with thirty subjects (23±4years), 18 female and 12 male. Each subject is instructed to grasp each object 10 times according to their convenience without any constraint in orientation and force of grasp. Hence for the 10 objects as mentioned earlier, grasped by 30 subjects grasping each one of them 10 times we get a total of 10×30×10(=3000) images for our experiment.

3. Methodology

The RGB images acquired through the sensor are first converted into gray level images.

3.1. Feature Extraction

Each image has certain features which can be used to distinctly identify it from any other image. Thus it is necessary to extract these features from the image for image matching as well as for classification. In previous works [10, 11, 12] classification has been performed using various different kinds of features. In the present work we have considered wavelet energies as features for both matching and classification purposes.

Wavelet transform is an efficient technique to extract features from any signal. Wavelet transform is based on small waves called wavelets which have varying frequency and a limited duration [13]. With respect to image processing, with the help of Wavelet transforms it is possible to represent images in multiple numbers of resolutions. This is advantageous because there might be some features that are undetected at a particular resolution but might be detected at some other resolution.

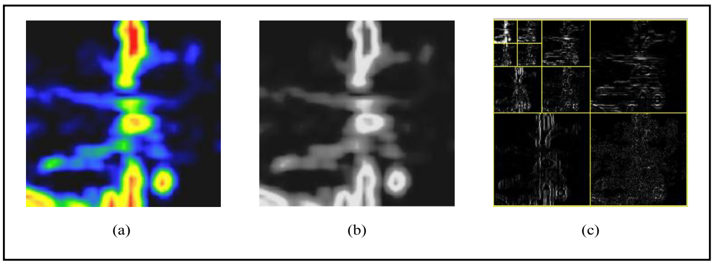

Figure 3. (a) Original RGB image, (b) Corresponding grayscale image, (c) Decomposition at level 3 for Haar Wavelet for (b)

The discrete wavelet transform (DWT) analyzes the images at different resolutions by decomposing the image into approximation and detail information. The outputs at the first decomposition level provides the detail D1 and approximation A1, respectively. The first approximation A1 is further decomposed into second approximation A2 and detail D2 and the process continued, until the desired result is obtained. At each level of decomposition, the detail coefficients include horizontal, vertical, and diagonal components. As images are 2-D arrays of gray level intensities, we work with 2-D DWT.

At ith level of decomposition, the original image will be represented by 3i+1 sub images, the approximation Ai and the details Hj, Vj and Dj for j=1 to i, where Hj denotes the horizontal detail, Vj denotes the vertical detail and Dj denotes the diagonal detail sub images [14]. The amplitudes in Hj, Vj and Dj correspond to high frequency (i.e. edges) in the respective directions. The wavelet energies for an M×N image for approximation, horizontal, vertical, and diagonal details at the ith level are respectively given by:

$$E_i^a=\Sigma_x\Sigma_y[A_i(x,y)]^2 \qquad \text{where } x=1 \text{ to } M/2^i \text{ and } y=1 \text{ to } N/2^i \tag{1}$$

$$E_i^h=\Sigma_x\Sigma_y[H_i(x,y)]^2 \qquad \text{where } x=1 \text{ to } M/2^i \text{ and } y=1 \text{ to } N/2^i \tag{2}$$

$$E_i^v=\Sigma_x\Sigma_y[V_i(x,y)]^2 \qquad \text{where } x=1 \text{ to } M/2^i \text{ and } y=1 \text{ to } N/2^i \tag{3}$$

$$E_i^d=\Sigma_x\Sigma_y[D_i(x,y)]^2 \qquad \text{where } x=1 \text{ to } M/2^i \text{ and } y=1 \text{ to } N/2^i \tag{4}$$

These energies represent the strength of the details of the image at the ith decomposition level. In the present work, the wavelet energies at three different decomposition levels are taken as image features. For each image, if we take 3rd level of decomposition, the energy estimated from the approximation coefficient has got a single component where as the energy estimated from each of the three detail coefficients consists of three components. Hence in this case the wavelet energy feature vector is a vector of dimensions 1×10 for each image. In our experiments there are 30 subjects and 10 objects with each subject grasping each object 10 times so we get a total of 3000 images and hence our feature space has dimensions of 3000×10.

In the present study, Daubechies mother wavelet of order 4 (db4), Symlet mother wavelet of order 4 (sym4), and Haar wavelet are used. After trials with the images, the features are selected from level 3 for classification and levels 3, 5 and 6 for matching. The decomposition level is chosen such that there is significant optimization between the loss of information and the classification or matching accuracy.

3.2. Classification

We have performed classification using three different classification techniques and then compared the classification accuracies and the run times thus obtained. In each case we have first divided the data in to a training set of data and a test set of data using cross-validation. In our work we have considered one-against-one (OVO) based classification approach. In one-against-all (OVA) approach, a binary classification process, there are only two possible classes; the features either belong to a particular class or they do not. However the OVO approach can perform multiclass classification.

3.2.1. k Nearest Neighbour (kNN) Classifier

In the k Nearest Neighbour classification algorithm, an object is put into the class which is the most common among its k nearest neighbours [15]. The value of k is that of a small positive integer. The distance between the neighbours is calculated using different methods, such as Euclidean distance, Mahalanobis distance, Manhattan distance etc. Experiments were conducted with different values of k ranging from 2 to 7 and it was found that k=3 gives the best results.

3.2.2. Linear Discriminant analysis (LDA) Classifier

Linear Discriminant analysis aims to find a linear combination of features that separates two or more classes [16]. The objective of LDA is maximizing the ratio of between-the-class to within-the-class distance thereby achieving the discrimination between classes.

3.2.3. Naïve Bayes Classifier

A Naïve Bayes Classifier is a probabilistic classifier based on Bayes theorem [17]. It assumes the independence of features within a class variable and uses the maximum likelihood principle for parameter estimation. This classifier is very robust to violations in its independence assumptions. It has produced efficient results in many applications including text classification [18] and medical diagnosis [15].

3.3. Image Matching

To establish the correspondence between existing images in a particular image database and a query image, an image matching technique is required for any kind of computer vision system [19],[20],[21].

For each subject, as we have 10 objects with each object being held 10 times, a feature space for 3rd level of wavelet decomposition, of dimensions 100×10 is obtained. For 30 subjects, 3000×10 feature space acts as the image database. Our goal is to provide a query image and see whether it is matched to a particular image in the image database of a particular object class. The query image will be represented by a 1×10 feature vector. If this 1×10 vector is similar to any of the vectors in the image database then a match occurs. Similarly for decomposition at level 5 and level 6 the feature spaces obtained are 3000×16 and 3000×19 respectively and the same process is followed for image matching. For matching, a metric in some form of similarity measure between the query image and the images in the database is required. There have been different kinds of similarity measures developed for image matching and retrieval over the years [22]. Examples include measures like Minowski form distance, Euclidean distance, quadratic form distance, Mahalanobis distance, Kullback-Leibler divergence etc. In the present work Euclidean distance has been used as a similarity measure. The Euclidean distance between the query image and all the images is computed and the class is determined from index corresponding to the minimum distance. Now the average Euclidean distance between the query image and all images belonging to each class is computed and the minimum average distance is computed to determine the class. If this class is equal to the class of the given query image as found before, then correct matching takes place.

4. Performance Analysis

The classification and matching results are provided in this section.

Table 1. Classification Accuracies for Haar Wavelet at Level 3

Classes	k-NN	LDA	Naïve Bayes
1	100%	100%	100%
2	80%	80%	60%
3	100%	80%	100%
4	60%	60%	60%
5	100%	100%	100%
6	80%	80%	60%
7	80%	80%	80%
8	60%	100%	80%
9	80%	60%	100%
10	100%	80%	60%
Mean	**84%**	82%	80%
Time (Sec)	0.0653	0.0398	0.0811

Table 2. Classification Accuracies for db4 Wavelet at Level 3

Classes	k-NN	LDA	Naïve Bayes
1	80%	80%	60%
2	80%	60%	80%
3	80%	60%	100%
4	100%	60%	40%
5	100%	80%	60%
6	80%	100%	80%
7	100%	80%	100%
8	60%	100%	60%
9	100%	60%	80%
10	80%	80%	60%
Mean	**86%**	76%	72%
Time (Sec)	0.0619	0.0402	0.0743

Table 3. Classification Accuracies for sym4 Wavelet at Level 3

Classes	k-NN	LDA	Naïve Bayes
1	80%	100%	80%
2	100%	60%	80%
3	60%	60%	80%
4	80%	40%	60%
5	80%	80%	80%
6	100%	80%	60%
7	80%	60%	100%
8	60%	60%	80%
9	100%	100%	80%
10	80%	80%	60%
Mean	82%	72%	76%
Time (Sec)	0.0597	0.0339	0.0716

Table 4. Matching Results

Wavelet	Decomposition Level	Matching Rate (%)	Time (Sec)
Haar	Level-3	96	0.0041
	Level-5	96	0.0058
	Level-6	88	0.0067
db4	Level-3	72	0.0042
	Level-5	76	0.0066
	Level-6	60	0.0070
sym4	Level-3	76	0.0043
	Level-5	88	0.0056
	Level-6	68	0.0066

4.1. Classification Results

Object shape classification is performed utilizing three classifier algorithms in one-against-one approach, i.e. multiclass classification for the 10 objects forming 10 distinct classes. The classification accuracies for one-versus-one (O-V-O) approaches have been shown for 3 different mother wavelets: Haar, Daubechies wavelet of order 4 (db4), Symlet wavelet of order 4 (sym4), each decomposed at level 3 in Tables 1, 2 and 3 respectively.

For Haar mother wavelet, the mean classification accuracy over all classes and over all classifiers is found to be 82%. k-NN provides the best performance with a mean classification accuracy of 84%. For Daubechies (db) mother wavelet of order 4, the mean classification accuracy over all classes and over all classifiers is found to be 78%. k-NN provides the best performance with a mean classification accuracy of 86%. For Symlet (sym) mother wavelet of order 4, the mean classification accuracy over all classes and over all classifiers is found to be 77%. k-NN provides the best performance with a mean classification accuracy of 82%.

4.2. Matching Results

For matching, experiments have been performed with 3 mother wavelets as stated before and each with 3 levels of decomposition: level 3, level 5 and level 6, as depicted in Table 4. Haar mother wavelet provides the highest matching rate of 96% at level 3 and level 5, with minimum run time of 0.0041sec in level 3. The next best matching rate of 88% is obtained for Haar mother wavelet at level 3 and sym4 wavelet at level 5.

5. Conclusion

The present work is concerned with finding the performances for the matching of a query image to one of the images in an existing image database as well as image classification based on wavelet energy features for object-shape recognition from tactile images. Experiments are performed using three mother wavelets: Haar, Daubechies wavelet of order 4 (db4) and Symlet wavelet of order 4 (sym4) and the wavelet energies are considered as features. k-NN classifier yields the best results with mean accuracy of 84% for all three mother wavelets. Best matching rate is found for Haar mother wavelet with mean matching rate of 93.33% over all three levels of decomposition. Comparing the results of image classification and matching, it is found that a maximum matching rate as high as 96% is obtainable with a computation time of 0.0041 sec using Haar mother wavelet at the 3rd level of decomposition; however, the maximum mean classification accuracy is considerably smaller, 86% for k-NN classifier (k=3) for db4 wavelet, with a greater computation time of 0.0619 sec. As is evident from theory as well as the experimental results, classification takes larger computation time, in this case, the minimum computation time for classification is about 8 times greater than that for matching using Euclidean distance as the similarity measure.

In future we aim to perform image matching using different forms of distance measures and compare their matching rates. We also aim to develop real time systems capable of image classification and matching. From the results obtained, it is only natural that image matching turns out to be a suitable candidate for real time applications taking up less computation time while providing significant performance improvement at the same time.

Acknowledgements

This study has been supported by University Grants Commission, India, University of Potential Excellence Program (UGC-UPE) (Phase II) in Cognitive Science, Jadavpur University and Council of Scientific and Industrial Research (CSIR), India.

References

[1] Khasnobish A., Jati A., Singh G.,Bhattacharyya S., Konar A.,Tibarewala D.N, et al., " Object Shape Recognition from Tactile Images using a Feed Forward Neural Network", IEEE Int. Joint Conf. Neural Networks, Brisbane, pp. 1-8, 2012.

[2] G.Canepa, M.Morabito, D.De Rossi, A.Caiti, T.Parisini, “Shape from touch by Neural Net”, Proceedings IEEE International Conference on Robotics and Automation, 1992
[3] Singh G., Jati A., Khasnobish A., Bhattacharyya S., Konar A., Tibarewala D.N., et al., “Object shape recognition from tactile images using regional descriptors”, IEEE 4th World Congress on Nature and Biologically Inspired Computing, 5-9Nov 2012, Mexico City, pp. 53-8.
[4] Singh G., Khasnobish A., Jati A., Bhattacharyya S., Konar A., Tibarewala D.N., et al. “Object-shape classification and reconstruction from tactile images using image gradient”, IEEE 3rd Int. Conf. Emerging Applications of Information Technology, 30Nov-1Dec 2012, Kolkata, India, pp. 93-6.
[5] Y. Yamada, et al., "A method of 3D object reconstruction by fusing vision with touch using internal models with global and local deformations," in Robotics and Automation, 1993. Proceedings., 1993 IEEE International Conference on, 1993, pp. 782-87 vol.2.
[6] B. Lazzerini and F. Marcelloni, "A fuzzy approach to 2D-shape recognition," Fuzzy Systems, IEEE Transactions on, vol. 9, pp. 5-16, 2001
[7] C. Jinhai and L. Zhi-Qiang, "Hidden Markov models with spectral features for 2D shape recognition," Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 23, pp. 1454-58, 2001
[8] Software for TactArray, Operator’s Guide, 2/1/2012
[9] Dahiya R.S, Valle M., “Tactile Sensing For Robotic Applications”, Sensors, Focus on Tactile, Force and Stress Sensors,I-Tech, Vienna, pp.289-304, 2008
[10] Haralick, R. M., Shanmugam, K., & Dinstein, I. H. (1973). Textural features for image classification. Systems, Man and Cybernetics, IEEE Transactions on, (6), 610-21.
[11] Cheng, Y. C., & Chen, S. Y. (2003). Image classification using color, texture and regions. Image and Vision Computing, 21(9), 759-76.
[12]Van de Wouwer, G., Scheunders, P., & Van Dyck, D. (1999). Statistical texture characterization from discrete wavelet representations. Image Processing, IEEE Transactions on, 8(4), 592-98.
[13] Gonzalez R.C., Woods R.E. Digital Image Processing 2nd ed., Prentice Hall, 2002
[14]Wu, X. Q., Wang, K. Q., & Zhang, D. (2005). Wavelet energy feature extraction and matching for palmprint recognition. Journal of Computer Science and Technology, 20(3), 411-18.
[15] Tom M. Mitchell., “Machine Learning”, McGraw Hill, 1997
[16] K. Etemad and R. Chellappa, "Discriminant analysis for recognition of human face images," in Audio- and Video-based Biometric Person Authentication. vol. 1206, J. Bigün, et al., Eds., ed: Springer Berlin Heidelberg, 1997, pp. 125-42
[17] Rish I.,“ An Empirical Study of the Naïve Bayes Classifier”, Int. workshop on Empirical Methods in AI, 2001.
[18] Kim, S. B., Han, K. S., Rim, H. C., & Myaeng, S. H. (2006). Some effective techniques for naïve bayes text classification. Knowledge and Data Engineering, IEEE Transactions on, 18(11), 1457-66.
[19] Van den Elsen, P. A., Pol, E. J., & Viergever, M. A. (1993). Medical image matching-a review with classification. Engineering in Medicine and Biology Magazine, IEEE, 12(1), 26-39.
[20] Haralock, R. M., & Shapiro, L. G. (1991). Computer and robot vision. Addison-Wesley Longman Publishing Co., Inc.
[21] Mahalakshmi, T., Muthaiah, R., & Swaminathan, P. Review Article: An Overview of Template Matching Technique in Image Processing. Research Journal of Applied Sciences, 4.
[22] Long, F., Zhang, H., & Feng, D. D. (2003). Fundamentals of content-based image retrieval. In Multimedia Information. Retrieval and Management (pp. 1-26). Springer Berlin Heidelberg.