Evaluation of the informativeness of multi-order image transforms

Lior Shamir* Nikita Orlov Ilya G. Goldberg

Laboratory of Genetics, NIA/NIH
251 Bayview Boulevard, Baltimore, MD 21224, USA
email: shamirl@mail.nih.gov

Abstract

We studied the informativeness of image features extracted from different lengths of image transform chains for the purpose of image classification. Image features were extracted from the raw images, image transforms, and second, third and fourth order of compound image transforms. The transforms used in this study are Fourier, Chebyshev, and Wavelet (symlet 5) transform. Experimental results show that image features extracted from first and second order of compound image transforms can in some cases be more informative than the image features extracted from the raw pixels, and can significantly contribute to the classification accuracy. However, chains of transforms longer than two do not improve the classification accuracy of the image datasets used in this study.

1. Introduction

Image features are descriptors that reflect the content of a given image. These can include high-frequency features such as textures, high-contrast features such as edges and shapes, polynomial decomposition of the image, statistical properties, color, and more. Large sets of low-level image features are commonly used for the purpose of general content-based image retrieval [10, 18, 13, 4].

Another increasingly popular application of computer vision which requires a large number of features is biomedical image analysis, in which the proliferation of imaging problems and classifiers to address them is acute [3, 2, 17, 21, 14]. The range of instrumentation and imaging modes available for capturing biological images multiplexed with the variety of morphologies exhibited by cells and tissues preclude a standard protocol for constructing problem-specific classifiers. The advent of High Content Screening (HCS) where the goal is to search through tens of thousands of images for a specific target morphology requires a flexible classification tool that allows any morphology to be used as a target. Since the variety of target morphologies is vast, a large set of image features is required to fully exploit the potential offered by HCS.

Present algorithms of image feature extraction do not capture all possible information that can be extracted from an image. Therefore, alternative representation of the images in the form of image transforms can provide additional information [19, 7, 16]. Murphy *et al.* [14] used Zernike moments [23] computed on the Fourier transform of an image to classify microscopy images of sub-cellular organelles. Hsu and Tseng [11] showed how image features extracted from wavelet transforms can be used for the recognition of ground features. Jiang and Yan [12] measured features from the Fourier transform to find coding regions (exons) in DNA sequences. [24] used simple features extracted from wavelet transforms to improve the classification accuracy of scenery images. Tsai [1] showed how 2D shapes can be recognized using spectral features extracted from the Fourier transform.

While features extracted from image transforms can be informative, these features can also be informative if extracted from chains of transforms (e.g., the Fourier transform of the Chebyshev transform of the image). Here we study the usefulness of using image features extracted from different lengths of compound image transforms for the purpose of image classification. In Section 2 we describe the classification method and the compound image transforms, and in Section 3 the experimental results are discussed.

2. Classification method

The image analysis method used in this study is *wnd-charm* [16, 21], which has been found effective for a wide range of image classification problems. *Wnd-charm* is an open source multi-purpose image classification tool that makes use of a large set of image features, extracted from several image transforms and compound transforms.

The image features used by *wnd-charm* cover high-contrast features (e.g., edges), textures (e.g., Haralick, Tamura), pixel statistics (e.g., multi-scale histogram, first four moments), and polynomial decomposition of the image. These include the following algorithms, which are described more thoroughly in [16, 21]:

1. **Zernike features** [23] are the absolute values of the coefficients of the Zernike polynomial approximation of the image as described by Murphy *et al*. [14], providing 72 image content descriptors.
2. **Multi-scale Histograms** computed using various number of bins (3, 5, 7, and 9), as proposed by Hadjidementriou *et al*. [8], providing 3+5+7+9=24 image content descriptors.
3. **First Four Moments** of mean, standard deviation, skewness, and kurtosis computed on image "stripes" in four different directions (0, 45, 90, 135 degrees). Each set of stripes is then sampled into a 3-bin histogram, providing $4\times4\times3$=48 image descriptors.
4. **Tamura Texture features** [22] of *contrast*, *directionality* and *coarseness* , such that the coarseness descriptors are its sum and its 3-bin histogram, providing 1+1+1+3=6 image features.
5. **Haralick features** [9] computed on the image's co-occurrence matrix as described in [14], and contribute 28 image descriptor values.
6. **Chebyshev Statistics** [6] - A 32-bin histogram of a 1×400 vector produced by Chebyshev transform of the image with order of N=20.

The image transforms that are used in this study are Fourier transform, Chebyshev transform, and Wavelet (symlet 5, level 1) transform. A detailed description of the implementation of these image transforms in *wnd-charm* can be found in [16, 21]. After the image features are computed, each feature is assigned with a weight based on its informativeness using Fisher Scores. The classification is based on a simple Weighted Nearest Neighbor rule, such that the Fisher scores are used as weights. A detailed description and performance analysis of the method is available in [16].

While the major downside of *wnd-charm* is its computational complexity, its large set of image content descriptors allows it to apply a systematic search for the most informative image features and image transforms. This is done by assigning a Fisher score to each image feature, so that the different features are weighted by their informativeness. When a feature vector of a test image needs to be classified, a simple Weighted Nearest Neighbor rule is applied such that the Fisher scores are the feature weights. This simple classification method is described more thoroughly in [16, 21]. The source code of *wnd-charm* is publicly open, and can be downloaded at [21].

To compare the informativeness of the different combinations of image transforms, the same set of image features was extracted from the different chains of transforms. That is, the same set of image features described earlier in Section 2 was extracted from the raw pixels, the Fourier transforms, Chebyshev transform, wavelet transform, and the different chains of compound transforms. The compound transforms include all possible combinations of transforms up to a length of four, such that each two consecutive transforms are different. The different chains of compound transforms are described in Figure 1.

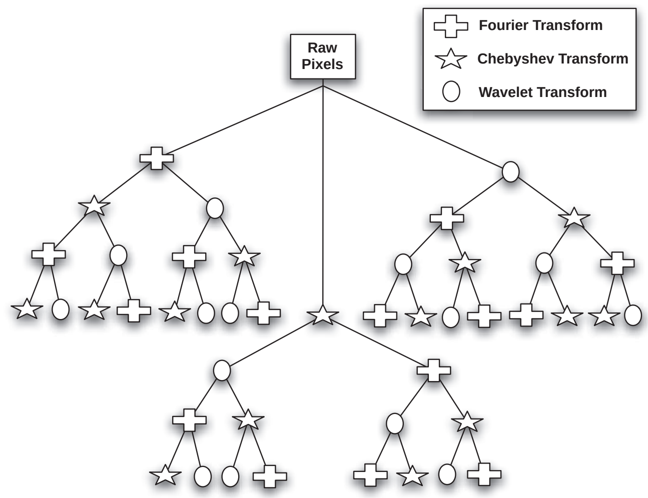

Figure 1. The chains of compound transforms

3. Experimental Results

The informativeness of the different chains of transforms was evaluated using several different image datasets. These datasets include biological images taken from IICBU 2008 benchmark suite [20], AT&T and Yale face datasets, Caltech 101 [5], and the COIL-20 object dataset [15]. The image datasets used for the experiment are listed in Table 1.

Dataset	No. of classes	No. of training images	No. of test images
Pollen	7	490	140
Hela	10	550	180
CHO	5	110	50
COIL-20	20	1000	400
AT&T	40	320	80
Yale	15	135	30
Caltech 101	101	6375	2125

Table 1. Datasets used for validation

To compare the informativeness of standard image features extracted from different lengths of chains of image transforms, we measured the classification accuracy of each dataset such that the image features were extracted only from chains of a certain length. The classification accuracies of the different datasets were measured by averaging 100 random splits to training and test sets, and the standard error in all cases is no larger than 0.45%. Figure 2 shows the classification accuracies of the different datasets using different lengths of chains of transforms. Clearly, the datasets included in the experiment have been studied before, and previously reported performance figures definitely outperform the figures reported in this paper. Here, however, these datasets are used only for the purpose of studying a more fundamental question, which is the correlation between the informativeness of the image features and the order of the image transform. Therefore, no attempt to propose a more efficient method for solving these specific dataset is made in this paper, and the paper does not discuss or propose such new methods.

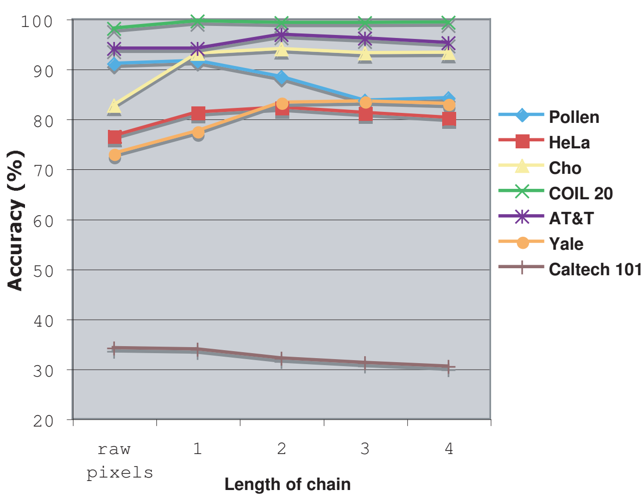

Figure 2. Classification accuracy using different lengths of chains of image transforms

As the figure shows, image features can in some cases be more informative than image features extracted from the raw pixels. However, for the experiments reported in this study, image features extracted from compound transforms do not become more informative when the chain of transforms is longer than two. In some datasets, such as Pollen and Caltech 101, image features become less informative as the chain of image transforms gets longer. However, for other datasets such as Yale, HeLa and CHO, the features extracted from the compound transforms are more informative than the features extracted from the raw pixels.

To increase image classification accuracy, image features extracted from image transforms can work in concert with image features extracted from the raw pixels. Thus, an image classifier can make use of features extracted from different chains of transforms to improve the overall classification accuracy. Since content descriptors extracted from transforms can reflect image content that is not captured by the features computed using the raw pixels, low-level image features extracted from image transforms can contribute to the overall classification accuracy of certain image classification problems. To test this contention, the classification accuracy was measured when all image features were computed using all chains of transforms shorter or equal to a certain length. For example, a chain of length zero included just image features extracted from the raw pixels, but a chain of length one used image features extracted from the Chebyshev, Fourier and Wavelet transform of the raw pixels in addition to the features computed using the raw pixels. Figure 3 shows the classification accuracy of the different image datasets using image features extracted from different lengths of chains of image transforms.

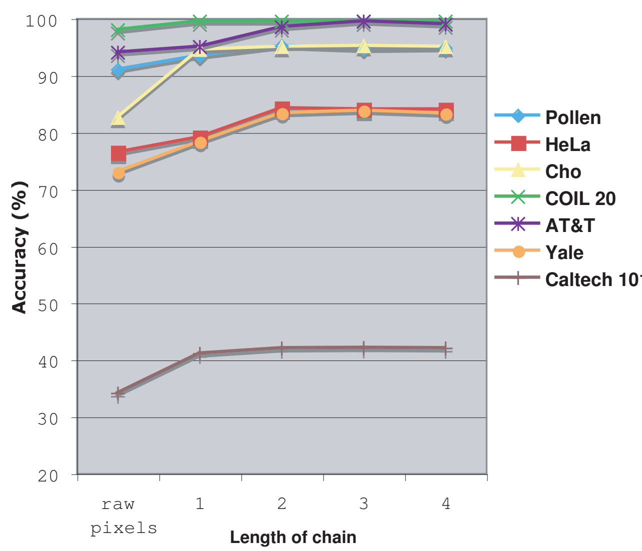

Figure 3. Classification accuracy using chains of transforms shorter or equal to a certain length

As the figure shows, low-level image features extracted from image transforms and compound image transforms can contribute to the overall classification accuracy. The figure also shows that chains of transforms longer than two generally do not improve the accuracy of the classification of the datasets tested in this study. Therefore, computing image content descriptors from long chains of transforms might consume significant computing resources, while contributing very little to the classification accuracy.

To estimate the informativeness of the different image features and the different transforms, we compared the accumulative Fisher scores (sum of the Fisher scores of the different bins) of the image features extracted from the different image transforms and compound transforms. For instance, the accumulative Fisher score of the Tamura features are the sum of the six content descriptors described in Section 2. Since the Fisher score reflects the discriminative power of each type of features, this analysis can allow comparing the discriminativeness of the different features in order to roughly estimate which features and chains of transforms are more informative. Figure 4 shows the Fisher scores of the different features and chains of transforms.

For the sake of convenience, and since the marginal contribution of chains longer than two was found to be low, the figure covers only features computed on chains of transforms that are not longer than two. As the figure shows, different features extracted from different combination of transforms can be informative for different datasets, and there are no combination of transforms that are clearly stronger than other combinations for all datasets. However, some of the features consistently perform better when extracted from a certain chain of transforms. A clear example is the Zernike features extracted from the Wavelet transform of the Fourier transform.

4. Conclusion

Here we showed that low-level image features can be informative if computed using image transforms and chains of compound image transforms, and these chains of transforms can contribute to the overall classification accuracy. Clearly, there are very many possible combinations of image transforms, multiplied by the different sets of image features that can be tested, and this work covers only a few of the possible experiments. However, this study suggests that low-level image features extracted from chains of image transforms can contribute to the performance of vision systems, and in some cases are more informative than the same set of features extracted from the raw pixels.

An obvious downside of computing image content descriptors from image transforms is the computational resources that should be sacrificed for computing the same features for each of the many possible combinations of transforms. Therefore, using this strategy might not be suitable in cases where response time is a primary concern.

While image features extracted from transforms and compound transforms can generally improve the accuracy of image classifiers, chains of transforms longer than two usually do not contribute to the overall classification accuracy. Therefore, in many cases the extraction of image features from longer chains of transforms might not justify the computational resources that should be sacrificed for that task.

5. Acknowledgments

This research was supported by the Intramural Research Program of the NIH, National Institute on Aging.

References

[1] J. Cai and Z. Q. Liu. Hidden markov models with spectral features for 2d shape recognitions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23:234–778, 2001. 1

[2] A. E. Carpenter. Software opens the door to quantitative imaging. *Nature Methods*, 4:120–121, 2007. 1

[3] A. E. Carpenter, T. R. Jones, M. R. Lamprecht, C. Clarke, I. H. Kang, O. Friman, D. A. Guertin, J. H. Chang, R. A. Lindquist, J. Moffat, P. Golland, and D. M. Sabatini. Cellprofiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biology*, 7:R100, 2006. 1

[4] T. Deselaers, D. Keysers, and H. Ney. Features for image retrieval: A quantitative comparison. *Proc. 26th Pattern Recognition DAGM Symposium*, pages 228–236, 2004. 1

[5] L. Fei-Fei, R. Fergus, and P. Perona. Learning generative visual models from few training examples: an incremental bayesian approach tested on 101 object categories. *Proc. Workshop Generative-Model Based Vision*, 2004. 2

[6] I. Gradshtein and I. Ryzhik. Table of integrals, series and products, 5th ed., 1994. Academic Press. 2

[7] I. B. Gurevich and I. V. Koryabkina. Comparative analysis and classification of features for image models. *Pattern Recognition and Image Analysis*, 16:265–297, 2006. 1

[8] E. Hadjidementriou, M. Grossberg, and S. Nayar. Spatial information in multiresolution histograms. *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, 1:702, 2001. 2

[9] R. M. Haralick, K. Shanmugam, and I. Dinstein. Textural features for image classification. *IEEE Trans. on Syst., Man and Cyber.*, 6:269–285, 1973. 2

[10] G. Heidmann. Unsupervised image categorization. *Image and Vision Computing*, 23:861–876, 2005. 1

[11] P. H. Hsu and Y. H. Tseng. Multiscale analysis of hyperspectral data using wavelets for spectral feature extraction. *Proc. 21st Asian Conference on Remote Sensing*, pages 821–827, 2000. 1

[12] R. Jiang and H. Yan. Segmentation of short human exons based on spectral features of double curves. *International Journal of Data Mining and Bioinformatics*, 2:15–35, 2008. 1

[13] R. Mao, Q. Iqbal, W. Liu, and D. P. Miranker. Case study: Distance-based image retrieval in the mobios dbms. *Proc. Intl. Conf. on Computer and Information Technology*, pages 49–57, 2005. 1

[14] R. F. Murphy, M. Velliste, J. Yao, and G. Porreca. Searching online journals for fluorescence microscopy images depicting protein subcellular location patterns. *Proc. IEEE Intl. Symp. on Bioinformatics and Biomedical Eng.*, pages 119–128, 2001. 1, 2

[15] S. A. Nene, S. K. Nayar, and H. Murase. Columbia object image library, 1996. Technical Report No. CUCS-006-96. Columbia University. 2

[16] N. Orlov, L. Shamir, T. Macura, J. Johnston, D. M. Eckely, and I. Goldberg. Wnd-charm: Multi-purpose image classification using compound image transforms. *Pattern Recognition Letters*, 29:1684–1693, 2008. 1, 2

[17] H. Peng. Bioimage informatics: a new area of engineering biology. *Bioinformatics*, 24:1827–1836, 2008. 1

[18] T. Quack, U. Monich, L. Thiele, and B. Manjunath. Cortina: a system for large-scale, content-based web image retrieval. *Proc. 12th ACM Intl. Conf. on Multimedia*, pages 508–511, 2004. 1

[19] K. Rodenacker and E. Bengtsson. A feature set for cytometry on digitized microscopic images. *Anal. Cell. Pathol.*, 25:1–36, 2003. 1

[20] L. Shamir, N. Orlov, D. M. Eckley, T. Macura, and I. G. Goldberg. Iicbu 2008 - a proposed benchmark suite for biological image analysis. *Medical & Biological Engineering & Computing*, 46:943–947, 2008. 2

[21] L. Shamir, N. Orlov, T. Macura, J. Johnston, D. M. Eckley, and I. G. Goldberg. Wndchrm - an open source utility for biological image analysis. *Open Source for Biology and Medicine*, 3:13, 2008. 1, 2

[22] H. Tamura, S. Mori, and T. Yamavaki. Textural features corresponding to visual perception. *IEEE Trans. on Syst., Man and Cyber.*, 8:460–472, 1978. 2

[23] M. R. Teague. Image analysis via the general theory of moments. *Journal of the Optical Society of America*, 70:920–930, 1979. 1, 2

[24] C. F. Tsai. Image mining by spectral features: A case study of scenery image classification. *Pattern Recognition*, 32:135–142, 2007. 1

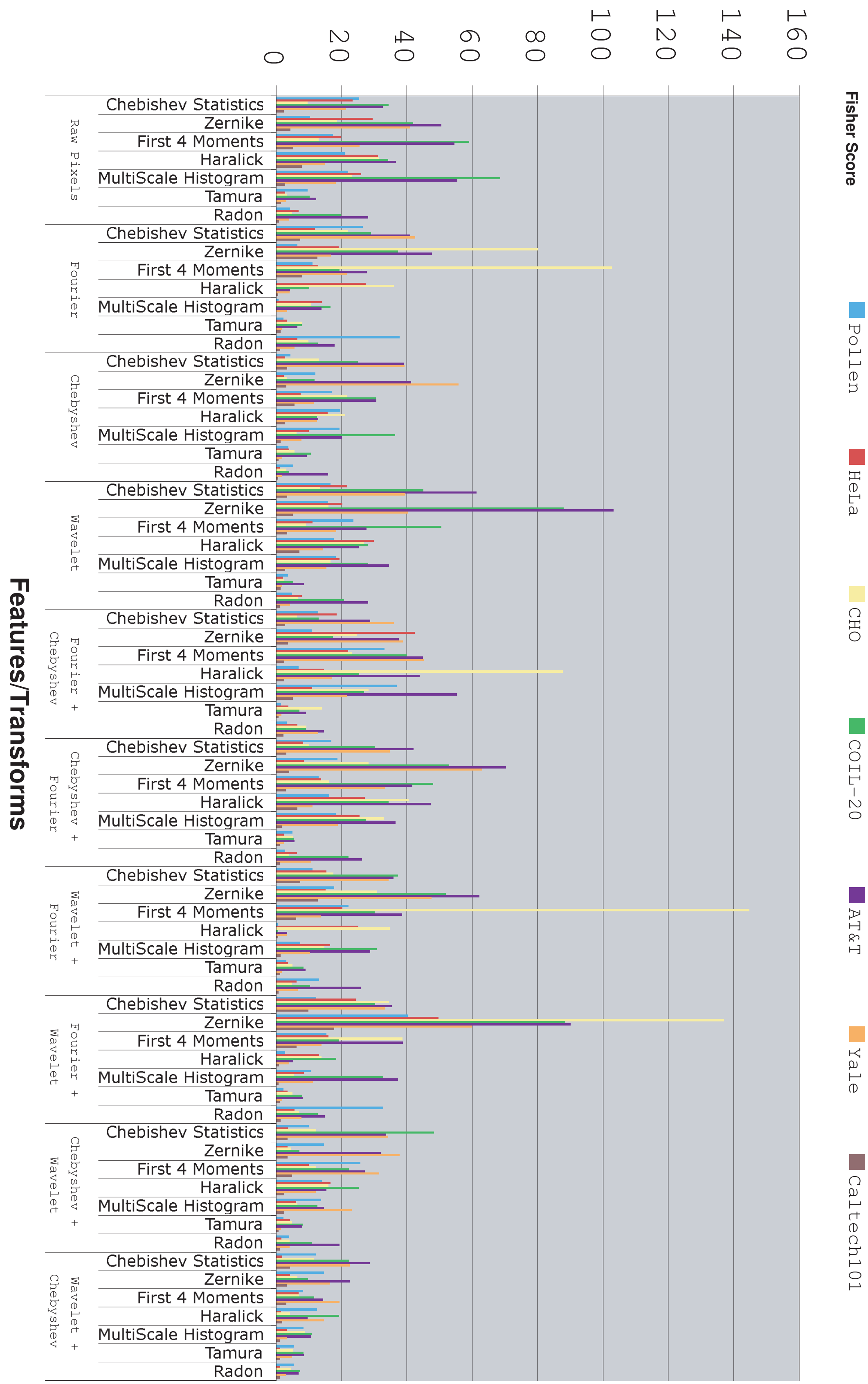

Figure 4. Fisher scores of the different features extracted from the different chains of transforms