



Published in final edited form as:

Mach Vis Appl. 2012 September 1; 23(5): 1047–1058. doi:10.1007/s00138-011-0349-5.

Improving class separability using extended pixel planes: a comparative study

Nikita V. Orlov, D. Mark Eckley, Lior Shamir, and Ilya G. Goldberg

National Institute on Aging /National Institutes of Health 251 Bayview Blvd, Bayview Research Center Bld, Suite 100, Baltimore, MD 21224, U.S.

Abstract

In this work we explored class separability in feature spaces built on extended representations of pixel planes (EPP) produced using scale pyramid, subband pyramid, and image transforms. The image transforms included Chebyshev, Fourier, wavelets, gradient and Laplacian; we also utilized transform combinations, including Fourier, Chebyshev and wavelets of the gradient transform, as well as Fourier of the Laplacian transform. We demonstrate that all three types of EPP promote class separation. We also explored the effect of EPP on suboptimal feature libraries, using only textural features in one case and only Haralick features in another. The effect of EPP was especially clear for these suboptimal libraries, where the transform-based representations were found to increase separability to a greater extent than scale or subband pyramids. EPP can be particularly useful in new applications where optimal features have not yet been developed.

Keywords

multi-scale representations; transforms; spectral features

1. Introduction

Several influential works have established the extended representation of image pixels as an approach to better classification. Tanimoto and Pavlidis [1] introduced scale pyramids as a tool for image processing, and this technique was explored further by Bourbakis and Klinger [2], and Burt and Adelson [3]. Other methods provided hierarchical representations of pixel planes: Subband transforms [4] by Simoncelli and Adelson, Scale-space method [5] by Lindeberg, and Steerable pyramids [6] by Freeman and Adelson. Subband transforms and the Laplacian pyramids [3] provide yet another type of pyramid-based EPP. We used Fourier, wavelets, gradient, Laplacian, and Chebyshev transforms to test other representations of pixels. In contrast to pyramids, transforms generate pixel planes with pixel locations redistributed in a non-linear manner, resulting in a variety of spatial patterns.

General multi-purpose feature libraries (MPF) are a subject of interest. Traditionally, for detecting an object within a given image, libraries of local features [7, 8] have been extensively used. A limitation of local features is that they operate on local pixel neighborhoods and tend to become over tuned to details of the target object, making generalizations more difficult [9]. Further, it has remained common practice to tailor feature libraries to reflect a priori knowledge of the given task, using manual selection, parameter tuning, or custom development. The broadening of application domains, particularly

biological [10-14] and biomedical [15-19] problems, has increased demand for automatic analysis and generalized image feature libraries, because the development of custom features in this vast problem space is cost-prohibitive [20]. Therefore, MPF libraries have begun to see broader use in projects concerned with expanded or cross-domain applications [21, 22]. This new trend in generalization of feature libraries includes neighborhood-based general-purpose libraries [23-26] as well as mixtures of general-purpose and task-specific libraries [27]. In this work we compare the relative contributions of EPP to classification, so we use a single set of general feature extraction algorithms (referred to as MPF subsequently) which includes a broad range of commonly used, general purpose algorithms for extracting numerical descriptors from image data. The specifics of this MPF library are described elsewhere [22].

We compare the separability in feature spaces computed from several types of EPP. In our approach, transforms are not used to produce features directly (e.g. as in the form of coefficients). Instead, they are used exclusively to generate a new pixel plane representing the contents of the original image in a different form. Due to the general nature of the feature algorithms in the MPF, these derived pixel planes (the EPP) can be processed by the MPF in the same way as the original pixel plane. Although a given algorithm in the MPF (e.g. Haralick textures) reports the same characteristic of the input image, its input images are different for each EPP, and thus the meaning of this feature is different with regards to what it says about the original image content. Thus, our method requires two-stages: a) generating EPP, and b) computing extended features using a single MPF library [13, 22].

Previously, we used this two-staged approach with a limited set of transforms and compared the performance of the resulting multi-purpose classifier to task-specific techniques. In this work, we evaluate the effect of upstream EPP using a total of nine derived pixel planes. In particular, we were interested to learn how transform-based EPP influences separability compared to scale and subband pyramids. We observed that EPP-based feature spaces significantly outperform those constructed without EPP. Particularly, we found that transform-based representations produce the most discriminative feature spaces. Additionally, we found that when a suboptimal feature library is incapable of providing separability on its own, the use of EPP can be essential for class separation.

2. Implementations Of EPP In Computational Chains

To generate EPP we map the raw pixel plane \mathcal{P} onto another plane \mathfrak{p} (or a set of planes $\{\mathfrak{p}^{(\alpha)}\}$), so that

$$\mathcal{P} \xrightarrow{\mathcal{F}} \bigcup_{\alpha} \mathfrak{p}^{(\alpha)}. \quad (2.1)$$

Implementation of EPP is based on the assumption that there exists a map \mathcal{F} for some given pixel plane \mathcal{P} . Should there exist an inverse map \mathcal{F}^{-1} , then we have isomorphism between the domain \mathcal{P} and its codomain \mathfrak{p} . Thus, the primacy of \mathcal{P} or \mathfrak{p} is uncertain, and spectral planes $\{\mathfrak{p}^{(\alpha)}\}$ may be treated as alternative definitions of the pixel plane \mathcal{P} (in Fig. 1a). These alternative definitions may improve separability of a feature space constructed from these codomains $\{\mathfrak{p}^{(\alpha)}\}$.

The second stage of our approach is computing extended features. The raw pixels are used to extract Numerical Patterns (NP, Fig. 1b). Ideally, any given imaging problem could be associated with its own unique set of filters, resulting in the most discriminative feature

space. We use a large library (\mathcal{L}) of multi-purpose features described in [22]. The projection of the pixel plane \mathcal{P} into feature space has the form

$$\vec{\psi} = \mathcal{L}(\mathcal{S}), \quad \mathcal{L}: \mathbb{R}^{M \times N} \xrightarrow{\mathcal{L}} \mathbb{R}^{L \times 1}, \quad (2.2)$$

where $\vec{\psi}$ is the feature vector, and numerical implementation of (2.2) defines a computational chain.

The majority of chains we used were computed on pixel planes produced with EPP, including multi-resolution analysis (or pyramids) and multi-transform analysis. The pyramids split the original pixel plane \mathcal{P} into a sequence of nested sub-planes $\{\mathbf{p}^{(\alpha)}\}$. For the Scale Pyramid (SP) [3] (Fig. 2a) we employed three levels of wavelet decomposition (symmetric family, Matlab wavelet toolbox2). The pyramids' indices α in (2.1) correspond to decomposition levels. The dimensions of the nested sub-pyramids, from level to level, were halved. Using (2.1)-(2.2) the SP feature vector can be written as

$$\vec{\psi}_{\text{SP}} = \bigcup_{\alpha} \mathcal{L}(\mathbf{p}^{(\alpha)}) = \bigcup_{\alpha} \vec{\psi}^{(\alpha)}. \quad (2.3)$$

The Subband Pyramid (SBP) is another type of multi-resolution representation used in this study. We describe SBP in classical terms of the discrete Fourier transform that maps the pixel plane \mathcal{P} onto a corresponding frequency domain $\mathbf{p} = \mathcal{F}[\mathcal{S}]$. Following [4] we define the SBP as a sequence of planes $\{\mathbf{p}^{(\alpha)}\}$ produced by bandpass frequency filters (indexing refers to different frequency bands). We constructed three 2D Finite Impulse Response filters on the corresponding rings in frequency domain. In our implementation (Fig. 2b) only three ($N_r = 3$) frequency bands were used; the SBP chain is shown in Fig. 3. The dimensions of all pixel planes are the same as that of the original image.

Pyramid transforms preserve image self-similarity in the sense that the original image is recognizable in the transformed images, which differ only in scale and/or resolution. Pyramids represent the original image as a set of pixel planes at different scales or resolutions, and thus do not spatially distort the original pixel plane (Fig. 2). Also, pyramids implement discrete downscale of the original pixel plane, it is a non-invertible transformation. In contrast, the result of an invertible transform $\mathbf{p} = \mathcal{F}[\mathcal{S}]$ (for instance the Fourier transform) the frequency domain can be legitimately used as an alternative representation of the original image. It preserves the image content while presenting a very different set of perceptual patterns to downstream feature extraction algorithms.

Feature extraction algorithms with known responses to image patterns (e.g. textures) can lose their meaning when applied to image transforms (e.g. textures in a Fourier-transformed image). This can hamper the perceptual interpretation of informative features selected by a classifier or dimensionality-reduction algorithm. In turn, this can reduce the utility of machine-based classification for certain applications, where understanding the source of the classification signal is important. On the other hand, many widely used image features (e.g. Zernike moments) are difficult to interpret perceptually even when applied to the original image pixels only. Strong correlations to easily interpreted perceptual patterns restrict the generality of feature extraction algorithms to those classification problems where such perceptual patterns exist. The use of these same algorithms on extended pixel planes extends their utility to new classification problems, though at the cost of their interpretability.

² <http://www.mathworks.com/access/helpdesk/help/toolbox/wavelet/>

Examples of such invertible transforms are given in Fig. 4 showing the Fourier (Fig. 4a, absolute value $|\mathcal{F}(\mathcal{S})|$ (0) is shown in logarithmic scale), Chebyshev (Fig. 4d, $\text{Ch}(\mathcal{S})$), and wavelet (Fig. 4e, level one, details) spatial transforms. Additionally we used non-invertible transforms: gradient (absolute value, $|\nabla S|$, Fig. 4b) and Laplacian¹ $\Delta\mathcal{S} = d^2\mathcal{S}/dx^2 + d^2\mathcal{S}/dy^2$ (Fig. 4c; not to be confused with the Laplace transform²). As one can see in Fig. 2, while SP and SBP produce pixel planes that are quite similar perceptually to the original pixels, transform-based EPP (Fig. 4) produce patterns completely different from those in the original image. We used this perceptual property as a guide for selecting the types of transforms and transform combinations for further study in this work. Thus, the five transform types were selected based on the diversity of the visual patterns produced relative to the original image, while limiting their number to a manageable set for testing transform combinations.

The set of patterns analyzed were further extended by using transforms sequentially (e.g. Fourier followed by Chebyshev). To limit computational time we tested four combinations that provided the greatest diversity of visual patterns: Fourier, Chebyshev and wavelets of the gradient (Fig. 4f, Fig. 4h, and Fig. 4i, respectively), and Fourier of the Laplacian (Fig. 4g), as shown in Table 1 (right column). We excluded some transform combinations – such as $\mathcal{W}(\text{Ch}[S])$, $|\nabla(\text{Ch}[S])|$, $|\nabla(\Delta[S])|$ or $\Delta(\text{Ch}[S])$ because they appeared to propagate noise, contributing little to discovering new patterns. The most important aspect of a transform is its ability to generate diverse perceptual patterns from any source image. Although our observations in this work and in previous work [22] indicate that the informativeness of features derived from transform combinations is highly dependent on the particular classification problem, there may be a theoretical basis for certain transform combinations being inherently more informative, or ones that are particularly prone to produce noise. We plan to undertake a more detailed exploration of these issues in the near future.

Fig. 5a gives a general scheme for computing spatial features (e.g. \mathcal{L} -chain based on (2.2), Fig. 5a), as well as implementing EPP with Image Transform Filters (ITF, Fig. 5b-5c). The output of ITF1 (Fig. 5b) consists of features computed on the original image and the five transforms described above. The output of ITF2 (Fig. 5c) consists of features from the original image and the four transform combinations. ITF3 combines the outputs of ITF1 and ITF2.

Somewhat similar approach by Lazebnik et al [28] is Spatial Pyramid Matching, where the image is split into a series of tiles (e.g. image sub-blocks) and the histograms of local (neighborhood-based) patches serve as features computed on the constituent tiles. Although it is possible to implement the MPF libraries for Spatial Pyramids in a similar fashion, we did not include it in our analysis because of the difficulty in making direct comparisons between an approach that sub-divides the image into tiles, and the other three implementations we tested where the pixel planes are indivisible.

3. Pattern Classifiers And Data Used In Experiments

We used three different classification methods: Weighted Neighbor Distance (WND) classifier, Support Vector Machines (SVM), and Radial Basis Functions networks (RBF). Having different classifiers report separability in feature spaces allows assessing properties of this space objectively while ensuring that we are measuring the effects of EPP rather than peculiarities of a particular classifier. The WND [22] algorithm relies on sample-to-class

¹Laplacian: <http://mathworld.wolfram.Laplacian.html>

²Laplace transform: <http://mathworld.wolfram.LaplaceTransform.html>

distances in a weighted space $\|\vec{w} \cdot (\vec{t} - \vec{X}_{cj})\|_2$, where \vec{t} , \vec{X}_{cj} are the test sample and its j -th vector to the training sample of class c , respectively. The weights \vec{w} are intended to reward/penalize features according to their discriminative power. They are multi-class Fisher scores [29], i.e., ratios of between-class and within-class variance for each feature.

Fisher scores were used for all three classifiers to rank features. When determining feature relevance, we selected top-scoring F features from the rank-ordered list for use in training. We used a greedy search where we increased the number of features used until the classification score stopped improving. The SVM classifier first performs a nonlinear mapping of data in high dimensional space and then constructs a separating hyperplane while maximizing the margin between the classes [30]. The SVM classifier uses an RBF kernel to compute projections into high-dimensional space. Finally, RBF networks [31] have no hidden layers and a direct procedure to compute the node weights using the Gaussian form of radial functions. While the feature rank was used in SVM and RBF for dimensionality reduction, WND used the Fisher scores as weights as well as for eliminating low-scoring features.

The total length (L) of the basic \mathcal{L} -chain is 376 entries. As one can see from the chain descriptions in Fig. 3, the sizes of the SBP and SP chains are $4L$ each. For transform-based chains described in Fig. 5, ITF1 is $6L$, ITF2 is $5L$, and the largest ITF3 is $10L$. The feature selection method described resulted in different numbers of features used for each classifier and each dataset used in testing. The range in the number of features used for the chains ITF3(\mathcal{L} ,3), ITF3(\mathcal{L} ,2), and ITF3(\mathcal{L}) were 120-338 (WND), 120-223 (RBF), and 78-137 (SVM).

Fisher scores are but one of the methods commonly used to rank features. In [15] we further explored different methods for selecting features and reducing dimensionality for a transform-based EPP scheme. Several feature ranking algorithms (Fisher weights, cosine similarity and Pearson correlation scores) were compared along with the subspace selection approach based on Maximal Relevance Minimum Redundancy [32]. We observed similar results for these different dimensionality reduction algorithms.

Nine data sets (see Table 2) were examined and grouped into three categories – biometric, biological and textural images. Several representative images from each dataset are shown in Fig. 6. The biometric category contains three face-recognition datasets: ATT ([33], Fig. 6a), Yale ([34], Fig. 6b) and Yale-B ([35], Fig. 6c). Image sizes were 112×92 pixels (ATT), 243×320 (Yale), and 350×350 (Yale-B, cropped). The biological group contains five datasets: Pollen ([36], Fig. 6d), CHO (fluorescent microscopy images of sub-cellular compartments for Chinese hamster ovary cells, [37]), HeLa ([10], Fig. 6e) and two biological datasets examining age-related muscle degeneration in the *C. elegans* pharynx terminal bulb (TB; [14, 38], Fig. 6f) and body wall muscle (BWM; [13], Fig. 6g & 6h). The TB was imaged with differential-interference contrast microscopy, while CHO, HeLa, and BWM used fluorescent microscopy with specific probes. The BWM images were pre-processed with a band-pass filter [13]. Images from these biological datasets vary in size: 25×25 (Pollen), 382×512 (CHO), and 382×382 (HeLa), 300×300 (TB) and 200×200 tiles for BWM. Finally, the commonly used Brodatz set [39] was used as the texture group. Brodatz images are 643×643 pixels. Each image was split into 16 tiles for training and testing since each texture type is represented with only a single image in this dataset.

All datasets contain multiple classes. Pollen, CHO, HeLa, BWM, and TB sets have seven, five, ten, four, and seven classes, respectively. Yale, Yale-B, and ATT face sets have 15, 10, and 40 classes, while the Brodatz texture set has 111 classes.

The reported classification results are based on cross-validation using ten random splits of the image data into training and test sets, with 80% of data being used for training and 20% for testing. The reported accuracy is an average for the ten independently trained classifiers in each test. Classification accuracy \mathcal{A} is defined as the ratio of test samples being called correctly versus total number of tested samples. In a classification problem where \mathcal{N}_C is the number of classes, the effective range for \mathcal{A} is $[1/\mathcal{N}_C, 1]$, making it difficult to directly compare classifier performance between problems with different \mathcal{N}_C . Thus, we propose an alternative measure of classifier performance that is independent of the number of classes, which we refer to as classification power (\mathfrak{P}) to distinguish it from conventional

$$\text{classification accuracy: } \mathfrak{P} = \begin{cases} (\mathcal{A} \mathcal{N}_C - 1) / (\mathcal{N}_C - 1), & \mathcal{A} > 1/\mathcal{N}_C \\ 0, & \mathcal{A} \leq 1/\mathcal{N}_C. \end{cases}$$

4. Results

To gain better understanding of the effect of EPP on separability of feature spaces, we performed a quantitative comparison of the relative separability properties of feature spaces calculated from subband pyramids, scale pyramids, and transforms. The results of each experiment are presented in tabular form, to allow inspection of the accuracy values for \mathcal{L} -space, pyramid feature spaces, and the different transform spaces (ITF1-ITF3).

Separability of the six spaces (\mathcal{L} , SBP, SP, ITF1, ITF2, and ITF3) was measured with three classifiers (WND, SVM, and RBF). Initially we used a complete feature library (\mathcal{L}) applied to the original pixel plane as the basis for comparison. The results of the first experiment are shown in Tables 3-5. We observed the \mathcal{L} -chain performed worst for all three classifiers, as expected, providing a baseline for minimum performance. In contrast, the ITF3-chain achieved the best results: the average (on all nine image datasets) classification power was 83.8% (WND), 84.4% (SVM), and 81.7% (RBF). Upgrading the basic \mathcal{L} -chain with transform-based EPP results in improved classification (for the ITF3-chain, an increase in power of 4.1% as average of the three classifiers, and up to 5.0% for RBF).

It is often the case that an available feature library is a poor choice for a given classification problem. We wanted to find out whether it is feasible to achieve an acceptable separability using a suboptimal library by compensating it with EPP. To answer this question, we made a subset of \mathcal{L} consisting only of texture features (\mathcal{L}_2 , consisting of Gabor, Tamura, and Haralick textures), and one consisting only of Haralick textures (\mathcal{L}_3). The \mathcal{L}_2 -chain has $L_2 = 62$ values, while the \mathcal{L}_3 -chain has $L_3 = 28$ values. The discrimination capacity of the \mathcal{L}_2 space declined severely (13.5% on average for all three classifiers, Tables 6-8) as compared with the original \mathcal{L} space, but less so for transform-based spectral spaces (on average, only 2% for ITF3(\mathcal{L}_2) against the best ITF3(\mathcal{L}) space). The most striking difference in separability was found between the best ITF3(\mathcal{L}_2) and the worst \mathcal{L}_2 space. This improvement in classification power (15.6% average for the three classifiers) is directly attributed to the increased separability provided by EPP.

The third library (\mathcal{L}_3 , including only Haralick features) (\mathcal{L}_3 -chain: $\mathcal{L}_3 = 28$ values) and subsequently calculated EPP. The basic \mathcal{L}_3 -chain, SBP(\mathcal{L}_3) and SP(\mathcal{L}_3) were completely ineffective, achieving close to zero classification power. The average power for those ineffective feature spaces did not exceed 22% (Tables 9-11). In contrast, all transform-based ITF(\mathcal{L}_3) spaces were reasonably separable. The power reported on ITF3(\mathcal{L}_3) was surprisingly close to the result of ITF3(\mathcal{L}) – the difference was 10%-11% (SVM and WND) and 19% (RBF). Therefore, the weakest \mathcal{L}_3 library demonstrated the biggest gain in separability, due to transform spaces (specifically, ITF3(\mathcal{L}) space).

5. Discussion

We expected EPP to generate separable feature spaces and the pyramids to be a comparable choice to transforms. EPP did perform well in all scenarios tested, but transform spaces were clearly and consistently better than pyramid spaces for classification. We also suspected the transforms to be inappropriate for some applications – such as the Pollen set (consisting of small images) or CHO (limited range of gray scale). Transforms were quite effective for all imaging problems we included in this study.

The results of experiments (Tables 3-11) clearly demonstrated the positive effect of EPP in improving separability. We also found that not all EPP were equally effective in detecting the discriminative content in the benchmark set. Based on the experiments conducted, we found that EPP's efficacy could be ordered from weakest to strongest as follows: SBP/SP, ITF1/ITF2, and ITF3. One heuristic explanation for this could be based on “degrees of freedom” of the pyramids as compared against the transforms. As agents of multi-scale approximation, pyramids obey the relations of self-similarity, either in scale (SP), or in frequency (SBP). Thus, their capacity to expand the pixel plane is limited to a particular axis of their selfsimilarity (e.g., scale or frequency). Comparisons of classification scores in the second and third experiments (Tables 6-11), illustrated the limited benefit of the pyramid-based representations. In contrast to the pyramids, the tested transforms were not limited by self-similarity. Transforms produce a broader range of perceptual patterns with no particular constraints in kind/dimension (Fig. 4), therefore promoting much improved use of the library. As we observed in the second and third experiments, the transformbased EPP generated the most separable spaces and produced reasonable results for all classification problems attempted, even when a suboptimal feature library was used (Tables 6-11).

The demonstrated advantage of transform-based EPP over pyramids bears further discussion. As noted in Sec.2, pyramids represent the original image as a set of pixel planes where, at each level in the pyramid, the pixel plane is similar in appearance to every other level as well as to the original image. In contrast, each transform-derived pixel plane has a very distinct appearance *totally unrelated* to the pixel plane from other transforms or to the original image pixel plane. A desired property of all general-purpose feature extraction libraries (as opposed to task-specific ones) is that they should be able to report similar patterns occurring at different scales and frequencies. However, this property leads to redundancy in the resulting features when only the scale of the image plane is varied, for example. Thus, the degree to which a feature library is scale and frequency invariant works directly against the variation in the image plane introduced by pyramids. An extreme case is when Haralick features are used alone (Tables 9-11). Though not strictly scale-invariant, varying only the scale or frequency of the pixel plane does not produce sufficiently different outputs for Haralick features to be effective for classification. In contrast, when the pixel planes are varied in regards to the types of textures they contain rather than just their scale (i.e. using transformbased EPP), Haralick features alone can classify the same image problems with marked improvement.

The use of image transforms as a source of features for pattern recognition is fairly common [37, 40] where generally, the coefficients of the transforms are used directly as image features. In contrast, here we use the transform coefficients as extended pixel planes (EPP) and process them using feature extraction algorithms that are more conventionally used on the original image pixels. By analogy, the use of transform coefficients as image features is similar in spirit to how original image pixels were used directly as image features in the past [41, 42]. Just as the use of feature extraction algorithms on raw image pixels has become more prevalent than using image pixels directly, here we show the relative benefits of using

feature extraction algorithms on image planes composed of transform coefficients rather than using the transform coefficients directly.

This comparison of different implementations of EPP was motivated by our interest in maximizing the diversity of patterns than can be analyzed given a fixed set of general feature extraction algorithms. In this work we did not seek to determine an optimal set of transforms or transform combinations. Indeed our experimental work here and elsewhere [22] indicates that the relative informativeness of the different transform-derived features is highly dependent on the specific imaging problem. Though this does not preclude identifying particularly informative EPPs (or particularly uninformative ones), it does indicate that there is probably not a universally optimal set. A fully automated feature reduction algorithm makes optimality of the transform set less important because it minimizes the impact of uninformative transforms.

In general, a search for separable feature space can be seen as fitting libraries (e.g. feature algorithms) to specifics of the given imaging problem. If class discrimination is weak, the library is probably a poor fit for the given problem. That implies that the library used is suboptimal for detecting discriminative content in pixel patterns for the task given. At the same time, for the very same task there may be other patterns that are more suitable for this library, – and these patterns could be found in pyramid and transform feature spaces, computed with the EPP approach. Situations where a library is a less than a perfect fit to the problem are very common. Hence, complementing difficult pixel patterns with rich alternative patterns promotes an increase in separability, as we observed in our second and the third experiments (Tables 6-11). A strategy of abstracting from the raw pixels to spectral spaces is a conceptual alternative to the search for a task-beneficial library. Spectral spaces help to address a common conceptual dilemma (fitting libraries to an imaging problem) and are the basis of a beneficial strategy – producing separable feature spaces.

The use of general-purpose libraries can be a good substitute for strategies of fitting features libraries to specific imaging problems by handpicking, custom development or tuning algorithms. Although these general-purpose libraries may be seen as weaker than task-specific libraries, the ideal fit of task-specific features rarely has a deep theoretical justification and is often overstated because it is empirically defined. From this standpoint, using multi-purpose libraries is often well justified both in terms of development cost and effectiveness for classification, as our experiments demonstrated.

We tested discriminating properties of the basic feature \mathcal{L} -library against the libraries based on EPP: frequency and scale pyramids, transforms and compound transforms. We demonstrated that EPP promotes class separation; improvement was detected in all three kinds of extended pixel-plane representations tested, two pyramids and transforms. The effect of EPP on weak feature libraries was especially clear for \mathcal{L}_2 and \mathcal{L}_3 chains; interestingly, only transform-based spectral spaces (ITF1, ITF2, ITF3) worked well compared with the basic \mathcal{L}_3 -chain and the pyramid filters. For example, one library consisting of only texture features (\mathcal{L}_2 -chain), had classification power improved by up to 25% from \mathcal{L}_2 to ITF3(\mathcal{L}_2).

We found the transform-based representations to have an advantage over the pyramid-based ones, which holds true in all experiments conducted. EPP could be particularly useful in a scenario when dealing with new applications where the best features have not been identified or are not developed yet.

Acknowledgments

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

References

1. Tanimoto S, Pavlidis T. A hierarchical data structure for picture processing. *Computer Graphics and Image Processing*. 1975; 4:104–119.
2. Bourbakis NG, Klinger A. A hierarchical picture coding scheme. *Pattern Recognition*. 1989; 22:317–329.
3. Burt P, Adelson E. The Laplacian pyramid as a compact image code. *IEEE Transactions on Communications*. 1983; 31:532–540.
4. Simoncelli, E.; Adelson, EH. Subband transforms. In: Woods, J., editor. *Subband image coding*. Kluwer Academic Publishers; Norwell: 1991.
5. Lindeberg, T. *Scale-space theory in computer vision*. Kluwer Academic Publishers; Norwell, MA: 1994.
6. Freeman WH, Adelson EH. The design and use of steerable filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 1991; 13:891–906.
7. Lowe DG. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*. 2004; 60:91–110.
8. Rosenfeld A. From image analysis to computer vision: an annotated bibliography, 1955-1979. *Computer Vision and Image Understanding*. 2001; 84:298–324.
9. Murphy K, Torralba A, Eaton D, Freeman W. Object detection and localization using local and global features. *Lecture Notes in Computer Science*. 2006; 4170:382–400.
10. Boland MV, Murphy RF. A Neural Network Classifier Capable of Recognizing the Patterns of all Major Subcellular Structures in Fluorescence Microscope Images of HeLa Cells. *Bioinformatics*. 2001; 17:1213–1223. [PubMed: 11751230]
11. Murphy RF. Automated interpretation of protein subcellular location patterns: implications for early detection and assessment. *Annals of the New York Academy of Sciences*. 2004; 1020:124–131. [PubMed: 15208189]
12. Ranzato M, Taylor PE, House JM, Flagan RC, LeCun Y, Perona P. Automatic recognition of biological particles in microscopic images. *Pattern Recognition Letters*. 2007; 28:31–39.
13. Orlov, N.; Johnston, J.; Macura, T.; Wolkow, C.; Goldberg, I. *International Symposium on Biomedical Imaging: From Nano to Macro*. Arlington, VA: 2006. Pattern recognition approaches to compute image similarities: application to age related morphological change; p. 1152-1156.
14. Johnston JL, Iser WB, Chow DK, Goldberg IG, Wolkow CA. Quantitative image analysis reveals distinct structural transitions during aging in *caenorhabditis elegans* tissues. *PLOS One*. 2008; 3:e2821. [PubMed: 18665238]
15. Orlov NV, Chen WW, Eckley DM, Macura TJ, Shamir L, Jaffe ES, Goldberg IG. Automatic classification of lymphoma images with transform-based global features. *IEEE Trans. on Information Technology in Biomedicine*. 2010; 14:1003–1013.
16. Orlov, NV.; Eckley, DM.; Shamir, L.; Goldberg, IG. *Visualization, Imaging and Image Processing*. Palma de Mallorca, Spain: 2008. Machine vision for classifying biological and biomedical images.
17. Shamir L, Ling SM, Scott W, Orlov N, Macura T, Eckley DM, Ferrucci L, Goldberg IG. Knee X-ray image analysis method for automated detection of Osteoarthritis. *IEEE Transactions on Biomedical Engineering*. 2009; 56:407–415. 2/2009. [PubMed: 19342330]
18. Shamir L, Ling SM, Scott W, Hochberg M, Ferrucci L, Goldberg IG. Early Detection of Radiographic Knee Osteoarthritis Using Computer-aided Analysis. *Osteoarthritis and Cartilage*. 2009 in press.
19. Shamir L, Ling S, Rahimi S, Ferrucci L, Goldberg IG. Biometric identification using knee X-rays. *International Journal of Biometrics*. 2009; 1:365–370. [PubMed: 20046910]
20. Shamir L, Delaney JD, Orlov N, Eckley DM, Goldberg IG. Pattern recognition software and techniques for biological image analysis. *PLOS Computational Biology*. 2010; 6
21. Gurevich IB, Koryabkina IV. Comparative analysis and classification of features for image models. *Pattern Recognition and Image Analysis*. 2006; 16:265–297.
22. Orlov N, Shamir L, Macura T, Johnston J, Eckley DM, Goldberg IG. WND-CHARM: Multi-purpose image classification using compound image transforms. *Pattern Recognition Letters*. 2008; 29:1684–1693. [PubMed: 18958301]

23. Leung T, Malik J. Representing and recognizing the visual appearance of materials using three-dimensional textures. *International Journal of Computer Vision*. 2001; 43:29–44.
24. Crosier M, Griffin LD. Using basic image features for texture classification. *International Journal of Computer Vision*. 2010; 88:447–460.
25. Ojala T, Maenpää T, Pietikainen M. Multiresolution grayscale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2002; 24:971–987.
26. Varma M, Zisserman A. A statistical approach to material classification using image patch exemplars. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2009; 31:2032–2047. [PubMed: 19762929]
27. Rodenacker K, Bengtsson E. A feature set for cytometry on digitized microscopic images. *Analytic cellular pathology*. 2003; 25:1–36.
28. Lazebnik, S.; Schmid, C.; Ponce, J. *CVPR*. NY: 2006. Beyond bags of features: Spatial pyramid pooling for recognizing natural scene categories; p. 2169-2178.
29. Fukunaga, K. *Introduction to statistical pattern recognition*. 2 ed.. Academic Press; San Diego: 1990.
30. Vapnik, VN. *Statistical learning theory*. Wiley-Interscience; New York: 1998.
31. Duda, R.; Hart, P.; Stork, D. *Pattern classification*. 2 ed.. John Wiley & Sons; New York, NY: 2001.
32. Peng H, Long F, Ding C. Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Tr. on Pattern Analysis and Machine Intelligence*. 2005; 27:1226–1238.
33. Samaria, F.; Harter, A. *Second IEEE Workshop on Applications of Computer Vision*. Sarasota, FL: 1994. Parameterization of a stochastic model for human face identification; p. 138-142.
34. Belhumeur PN, Hespanha JP, Kriegman KJ. Eigenfaces vs. Fisher-face: recognition using class specific linear projection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 1997; 19:711–720.
35. Georgiades AS, Belhumeur PN, Kriegman DJ. From few to many: illumination cone models for face recognition under variable lighting and pose. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2001; 23:643–660.
36. Duller AWG, Duller GAT, France I, Lanb HF. A pollen image database for evaluation of automated identification systems. *Quaternary Newsletter*. 1999; 89:4–9.
37. Boland M, Markey M, Murphy R. Automated recognition of patterns characteristic of subcellular structures in fluorescence microscopy images. *Cytometry*. 1998; 33:366–375. [PubMed: 9822349]
38. Shamir L, Macura T, Orlov N, Eckley DM, Goldberg IG. ICBU 2008 - a proposed benchmark suite for biological image analysis. *Medical & Biological Engineering & Computing*. 2008; 46:943–947. [PubMed: 18668273]
39. Brodatz, P. *Textures*. Dover Pub.; New York, N.Y.: 1966.
40. Chebira A, Barbotin Y, Jackson C, Merryman T, Srinvasa G, Murphy RF, Kovacevic J. A multiresolution approach to automated classification of protein subcellular location images. *BMC Bioinformatics*. 2007; 8
41. Hu M-K. Visual pattern recognition by moment invariants. *IRE Transactions on information theory*. 1962; 8:179–187.
42. Pavlidis, T. *Algorithms for graphics and image processing*. Springer-Verlag; Berlin-Heidelberg-New York: 1982.

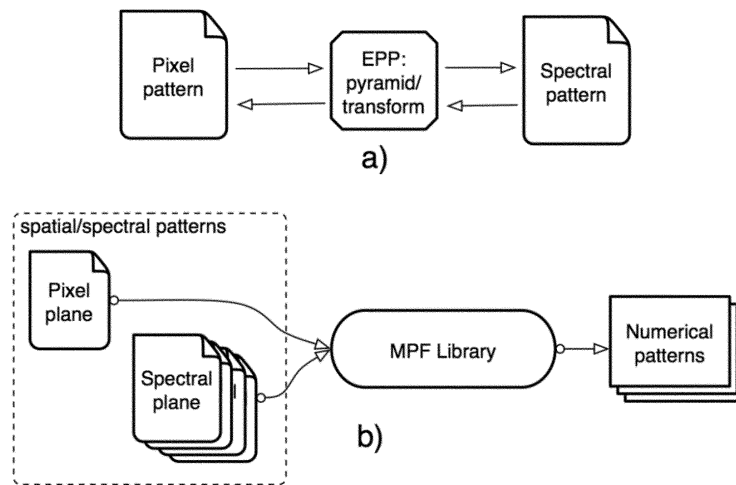


Fig. 1. Scheme of EPP and MPF library in the computational framework. a) EPP produces spectral planes with distinct visual patterns. b) MPF library is applied uniformly to the original pixel plane and to spectral planes generated by EPP. MPF library produces numerical values representing image content to be used by the statistical classifier.

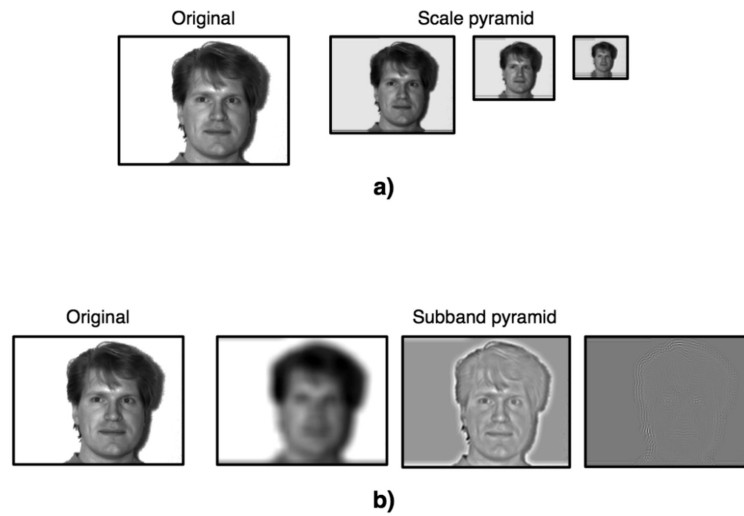


Fig. 2. Examples of EPP using Scale (a) and Subband (b) Pyramids. A Scale Pyramid produces a set of images where each level is scaled by half. A Subband Pyramid is implemented by applying a windowed bandpass filter to the original image resulting in a set of images where each subsequent level consists of higher frequency information.

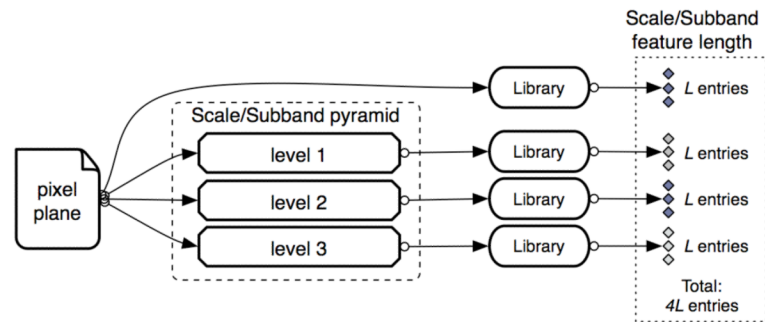


Fig. 3. The computational chain for features computed with MPF library on Scale/Subband pyramids. The MPF library is applied to the original image, resulting in a feature vector of length L . The same MPF is applied to each level of the pyramid to generate level-specific features, resulting in an overall feature vector length of $4L$.

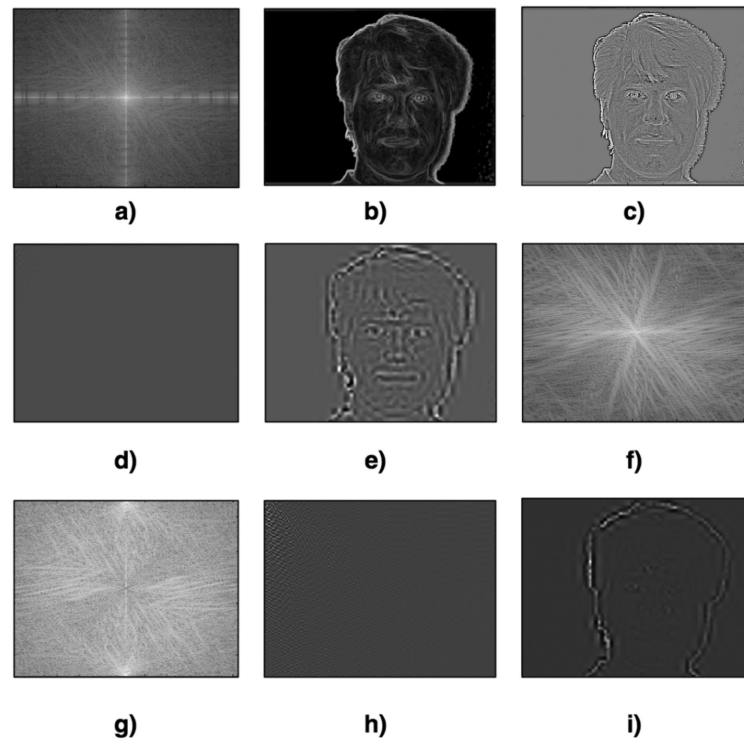


Fig. 4. Examples of transforms for one instance from the Yale dataset. Image transforms shown: a) $(\mathcal{C}[S])$, b) $|\nabla S|$, c) ΔS , d) $\text{Ch}[S]$, e) $\mathcal{W}[S]$, f) $\log(|\nabla S|)$, g) $\log(|\Delta S|)$, h) $\text{Ch}[|\nabla S|]$, and i) $\mathcal{W}[|\nabla S|]$. The visual patterns in pixel planes produced by these transforms are distinct from each other as well as from the original image.

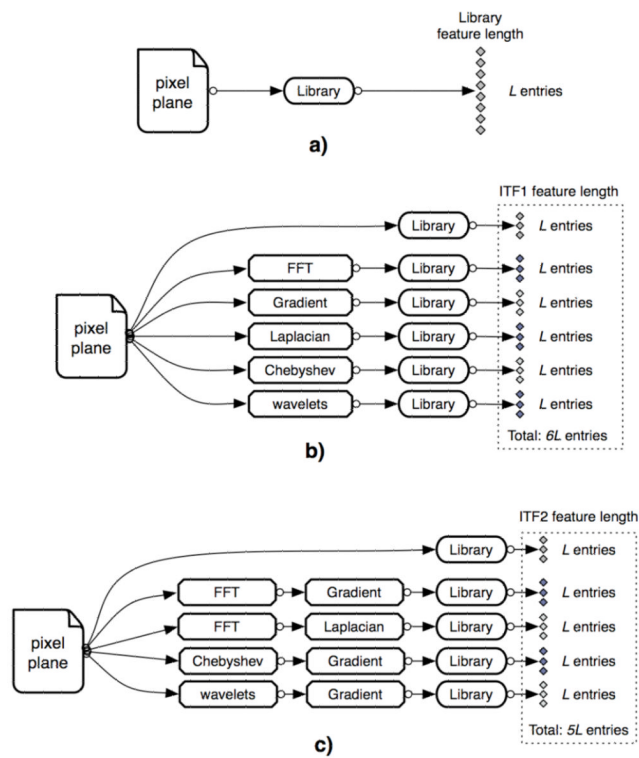


Fig. 5. Three chains computed on the original pixel plane (a: \mathcal{L} -chain), on planes produced by single transforms (b: ITF1(\mathcal{L}) chain), and by compound transforms (c: ITF2(\mathcal{L}) chain). The set of features computed from each transform-derived pixel plane is transform-specific so that additional transforms and transform combinations result in additional features in the final feature vector.

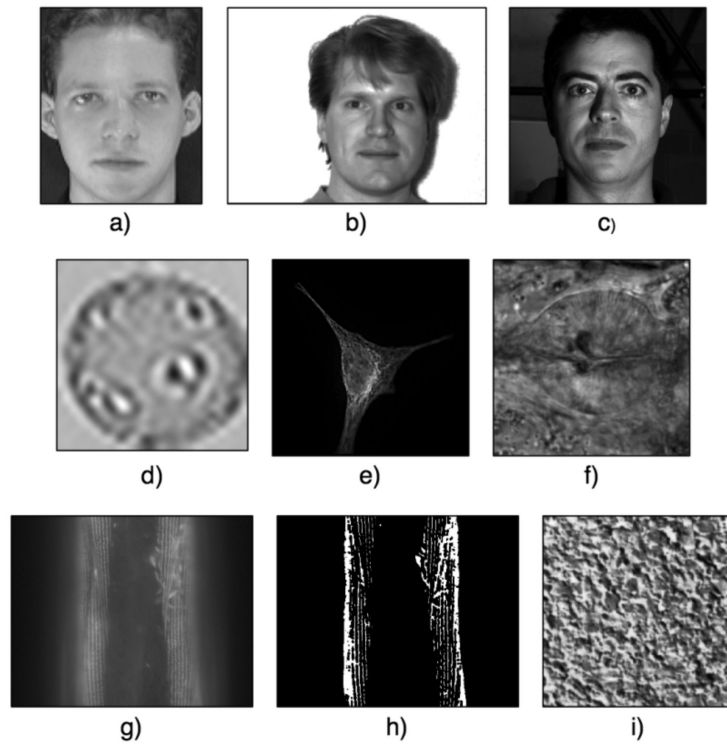


Fig. 6. Representative images from the benchmark data set used in this work: a) ATT set, b) Yale set, c) Yale-B set, d) Pollen set, e) HeLa set (microtubules), f) *C. elegans*, terminal bulb, g) *C. elegans*, body wall muscle (raw data), h) *C. elegans*, body wall muscle (processed: bandpass filter followed by thresholding), and i) Brodatz set.

Table 1

Image transforms and compound transforms used in ITF-chains.

One-level transforms	Compound transforms
$ \mathcal{F}(S) $	$ \mathcal{F}(\nabla S) $
$ \nabla S $	$ \mathcal{F}(\Delta S) $
ΔS	Ch ($ \nabla S $)
Ch (S)	$\mathcal{W}(\nabla S)$
$\mathcal{W}(S)$	

Table 2

Data sets used

Data set	Number of classes	Image size	Training set	Test set
Pollen	7	25×25	490	140
CHO	5	382×512	80	262
HeLa	10	382×382	590	272
<i>C.elegans</i> , BWM	4	200×200	740	495
<i>C.elegans</i> , TB	7	300×300	959	966
ATT	40	112×92	200	200
Yale	15	243×320	90	75
Yale-B	10	350×350	4680	1170
Brodatz	112	643×643	896	896

Comparison of separability properties (as classification power) for \mathcal{L} and spectral spaces using the WND classifier. Classification power (see text) is a percentage where 100% is perfect classification, and 0% is at the noise-floor for a particular classification problem (i.e. 100 / number of classes). This table illustrates that a modest improvement in classification can be achieved with EPP even when using an extensive feature library.

Table 3

Set	\mathcal{L}	SBP(\mathcal{L})	SP(\mathcal{L})	ITF1(\mathcal{L})	ITF2(\mathcal{L})	ITF3(\mathcal{L})
Pollen	94.3	93.9	94.7	94.1	92.8	95.8
CHO	95.5	95.7	96.1	97.6	96.2	96.8
HeLa	78.8	81.2	83.8	83.4	84.2	85.5
<i>C.elegans</i> , BWM	46.5	47.4	48.0	49.3	49.2	52.4
<i>C.elegans</i> , TB	43.9	46.1	48.5	50.4	47.0	51.7
ATT	99.3	99.7	99.7	98.7	99.1	99.5
Yale	81.3	85.5	83.9	80.0	81.2	80.2
Yale B	96.4	97.2	99.8	97.3	96.2	97.5
Brodatz	90.8	91.7	90.5	92.8	91.5	94.5
Average	80.8	82.0	82.8	82.6	81.9	83.8

Table 4

Same comparisons as in Table 3 using the SVM classifier, illustrating that the benefits of EPP are not classifier dependent.

Set	χ	SBP(%)	SP(%)	ITF1(%)	ITF2(%)	ITF3(%)
Pollen	93.9	96.5	96.8	96.5	94.5	96.0
CHO	95.0	96.0	97.3	96.3	97.8	97.6
HeLa	83.2	86.8	89.8	88.8	88.3	90.0
<i>C.elegans</i> , BWM	37.3	39.3	38.0	41.2	40.8	42.6
<i>C.elegans</i> , TB	44.1	45.8	48.5	51.1	47.9	53.1
ATT	98.3	98.8	99.2	99.4	98.9	99.7
Yale	81.4	83.5	84.7	83.7	86.0	86.0
Yale B	96.6	97.7	100	97.3	97.3	97.4
Brodatz	91.2	93.7	92.7	95.0	93.4	97.3
Average	80.1	82.0	83.0	83.3	82.8	84.4

Table 5

Same comparisons as in Table 3 using the RBF classifier, illustrating that the benefits of EPP are not classifier dependent.

Set	χ	SBP(%)	SP(%)	ITF1(%)	ITF2(%)	ITF3(%)
Pollen	89.5	89.6	91.2	93.8	90.6	93.1
CHO	88.8	77.8	87.2	95.1	95.7	95.2
HeLa	75.1	76.6	82.3	80.6	80.5	83.3
<i>C.elegans</i> , BWM	43.2	44.8	44.8	46.4	44.0	49.4
<i>C.elegans</i> , TB	39.6	39.3	40.9	42.2	40.0	43.7
ATT	99.6	98.7	99.6	97.5	99.3	99.1
Yale	75.4	82.2	79.9	81.7	78.1	80.8
Yale B	96.5	97.0	99.7	98.4	96.2	98.4
Brodatz	82.1	89.7	89.6	92.6	91.5	92.0
Average	76.6	77.3	79.5	80.9	79.5	81.7

Table 6

Comparison of the separability properties of a subset of MPF features containing only textures (\mathfrak{L}_2). Features resulting from the original image alone (\mathfrak{L}_2) are compared to five sets of extended pixel planes (SBP: Sub-band Pyramid, SP: Scale Pyramid, ITF1: single transforms, ITF2: compound transforms, ITF3: \mathfrak{L}_2 , ITF1 and ITF2 combined). The WND classifier was used in all tests. This table illustrates that a limited feature library can be improved by using EPP, especially transform-based EPP.

Set	\mathfrak{L}_2	SBP(\mathfrak{L}_2)	SP(\mathfrak{L}_2)	ITF1(\mathfrak{L}_2)	ITF2(\mathfrak{L}_2)	ITF3(\mathfrak{L}_2)
Pollen	86.4	91.9	91.6	93.3	90.0	92.4
CHO	83.1	91.6	87.2	97.0	91.8	97.0
HeLa	65.5	77.7	81.0	83.8	75.5	82.7
<i>C.elegans</i> .BWM	33.2	43.7	42.8	45.4	45.2	48.8
<i>C.elegans</i> .TB	37.3	35.2	39.9	50.5	40.0	49.3
ATT	93.6	94.9	95.8	98.2	96.0	98.5
Yale	68.2	72.6	79.5	78.3	77.5	80.8
Yale B	90.5	94.2	99.5	96.1	94.0	97.0
Brodatz	79.6	82.1	86.9	92.4	90.5	93.7
Average	70.8	76.0	78.2	81.7	77.8	82.2

Table 7

Same comparisons as Table 6 using the SVM classifier, illustrating that the effect of EPP on classification is not classifier-dependent.

Set	\mathcal{L}_2	SBP(\mathcal{L}_2)	SP(\mathcal{L}_2)	ITF1(\mathcal{L}_2)	ITF2(\mathcal{L}_2)	ITF3(\mathcal{L}_2)
Pollen	85.5	92.8	94.5	95.3	88.1	94.1
CHO	88.2	93.1	44.8	97.0	96.0	97.5
HeLa	77.5	82.8	83.7	89.5	80.1	90.4
<i>C.elegans</i> , BWM	31.3	37.0	32.6	37.3	36.6	40.6
<i>C.elegans</i> , TB	37.9	41.5	44.1	47.3	41.6	48.6
ATT	91.8	96.0	98.7	98.2	94.7	98.8
Yale	75.1	74.3	78.3	80.5	77.5	81.2
Yale B	92.7	95.3	99.5	97.6	91.4	97.2
Brodatz	89.9	87.6	91.0	94.5	93.0	95.5
Average	74.4	77.8	74.1	82.0	77.7	82.7

Table 8

Same comparisons as Table 6 using the RBF classifier, illustrating that the effect of EPP on classification is not classifier-dependent.

Set	\mathcal{X}_2	SBP(\mathcal{X}_2)	SP(\mathcal{X}_2)	ITF1(\mathcal{X}_2)	ITF2(\mathcal{X}_2)	ITF3(\mathcal{X}_2)
Pollen	77.6	81.4	77.6	86.7	81.4	88.3
CHO	77.0	89.5	84.7	94.7	87.5	95.5
HeLa	59.4	70.4	72.7	81.6	72.8	81.2
<i>C.elegans</i> , BWM	31.0	44.4	32.8	42.0	40.1	41.2
<i>C.elegans</i> , TB	31.1	29.3	27.2	41.7	32.2	41.0
ATT	74.4	91.3	92.4	94.8	84.7	95.2
Yale	18.8	66.1	67.9	74.6	74.1	83.7
Yale B	58.5	74.7	94.0	96.7	89.7	96.4
Brodatz	36.4	59.2	73.7	79.8	78.6	86.6
Average	51.6	67.4	69.2	77.0	71.2	78.8

Table 9

Comparison of the separability properties of a subset of MPF features containing only Haralick textures (\mathfrak{L}_3). Features resulting from the original image alone (\mathfrak{L}_3) are compared to five sets of extended pixel planes (SBP: Sub-band Pyramid, SP: Scale Pyramid, ITF1: single transforms, ITF2: compound transforms, ITF3: \mathfrak{L}_3 , ITF1 and ITF2 combined). The WND classifier was used in all tests. This table illustrates how a very poor feature library can be substantially improved by using EPP, especially transform-based EPP.

Set	\mathfrak{L}_3	SBP(\mathfrak{L}_3)	SP(\mathfrak{L}_3)	ITF1(\mathfrak{L}_3)	ITF2(\mathfrak{L}_3)	ITF3(\mathfrak{L}_3)
Pollen	0	0	0	75.8	39.8	78.1
CHO	0	29.2	33.5	97.1	86.7	96.0
HeLa	0	12.8	32.5	76.3	62.6	76.6
<i>C.elegans</i> , BWM	0	0	0	36.0	30.6	39.8
<i>C.elegans</i> , TB	0	0	0	28.8	19.5	32.8
ATT	0	0	15.8	84.9	35.1	83.3
Yale	0	18.3	11.0	77.8	71.3	76.3
Yale B	50.1	85.3	79.0	88.6	74.3	90.0
Brodatz	0	4.1	2.1	71.7	59.1	76.2
Average	5.6	16.6	19.3	70.8	53.2	72.1

Table 10

Same comparisons as Table 9 using the SVM classifier, illustrating that the effect of EPP on classification is not classifier-dependent.

Set	\mathcal{L}_3	SBP(\mathcal{L}_3)	SP(\mathcal{L}_3)	ITF1(\mathcal{L}_3)	ITF2(\mathcal{L}_3)	ITF3(\mathcal{L}_3)
Pollen	0	15.3	0	80.5	42.0	80.7
CHO	0	36.8	44.7	98.2	89.1	97.6
HeLa	0	13.3	35.3	85.6	73.4	85.8
<i>C.elegans</i> , BWM	0	0	0	33.2	27.7	35.7
<i>C.elegans</i> , TB	0	0	0	32.8	21.9	34.2
ATT	0	15.1	3.5	85.6	32.4	84.0
Yale	0	24.8	23.1	77.7	65.6	74.3
Yale B	50.6	83.3	75.0	90.8	63.3	89.7
Brodatz	0	5.1	2.1	84.7	67.2	86.3
Average	5.6	21.5	20.4	74.3	53.6	74.3

Table 11

Same comparisons as Table 9 using the RBF classifier, illustrating that the effect of EPP on classification is not classifier-dependent.

Set	\mathcal{L}_3	SBP(\mathcal{L}_3)	SP(\mathcal{L}_3)	ITF1(\mathcal{L}_3)	ITF2(\mathcal{L}_3)	ITF3(\mathcal{L}_3)
Pollen	0	0	0	64.4	39.3	65.9
CHO	0	0	27.2	90.3	77.1	92.0
HeLa	0	0	29.2	74.2	54.2	74.4
<i>C.elegans</i> , BWM	0	0	0	36.4	32.5	38.5
<i>C.elegans</i> , TB	0	0	0	28.7	19.5	33.3
ATT	0	0	0	65.1	23.6	61.6
Yale	0	10.0	0	69.2	56.7	70.1
Yale B	0	56.6	42.0	76.6	56.5	84.4
Brodatz	0	2.1	1.1	39.8	18.2	42.0
Average	0	7.6	11.1	60.5	42.0	62.5