MULTIRESOLUTION LINEAR DISCRIMINANT ANALYSIS: EFFICIENT EXTRACTION OF GEOMETRICAL STRUCTURES IN IMAGES

Sinisa Todorovic and Michael C. Nechyba

Dept. of Electrical and Computer Engineering, University of Florida
P.O. Box 116200, Gainesville, FL 32611-6200
{sinisha , nechyba}@mil.ufl.edu

ABSTRACT

Currently popular feature extraction tools (e.g., Gabor, wavelet analysis) do not economically represent edges in images. As a step towards solving this problem, the wedgelet transform was recently proposed [1]; this transform provides nearly optimal representation of objects in the Horizon model, as measured by the minimax mean-squared error. However, there is no reason to assume that the components useful for representing pixel values must also be useful for discriminating between regions in an image. Thus, having the successful extraction of edges as our goal, we propose a novel image analysis method—namely, multiresolution linear discriminant analysis (MLDA). In MLDA, analogously to the wedgelet transform, we seek directions that are efficient for discrimination. The MLDA framework comprises the following components: the MLDA atom, dictionary, tree, graph, and MLDA-based algorithms. In this paper, we explain these components and demonstrate the powerful expressiveness of MLDA, which gives rise to fast geometrical-structure-analysis algorithms.

1. INTRODUCTION

For years now, active research has been conducted in the area of wavelet-based image processing. However, recent findings on human vision and natural image statistics [2, 3] provide a host of arguments which seem to undermine the popularity of wavelets. It has been reported that cortical cells are not only highly sensitive to the location and scale, but also to the orientation and elongation of stimuli. Moreover, the basis elements which best "sparsify" natural scenes are highly direction-specific, unlike wavelets. Finally, it is well known that wavelets do not economically represent even straight edges, let alone more complicated geometrical structures in images. Therefore, all the aforementioned arguments suggest that there is a need for new image analysis methods that should exhibit, aside from the multiscale and localization properties of waveletes, also, characteristics that account for concepts beyond the wavelet framework.

Recently, the wedgelet transform has been proposed [1], as a step toward improved extraction and representation of edges in images. The wedgelet transform is a harmonic analysis method for nearly optimal representation of binary images consisting of piecewise constant regions separated by smooth boundaries (i.e. the Horizon model). A wedgelet is a piecewise constant function on either side of a line that intersects a dyadic square. The multiscale wedgelet representation of an image consists of a set of wedgelets supported by dyadic squares of varying sizes that partition the analyzed image. The wedgelet recursive partitioning is optimized over multiple criteria [1, 4] which essentially minimize the complexity-penalized mean-squared error. Clearly, the wedgelet representation seeks a projection that best represents an image in a least-squares sense. However, there is no reason to assume that the components useful for representing pixel values must also be useful for discriminating between regions in an image. Therefore, we propose a substantially different image analysis method which maximizes the distance between the means of the two regions while minimizing the variance within each region. Thus, we implement multiscale linear discriminant analysis (MLDA) which seeks for directions that are efficient for discrimination.

Despite its similarity to the wedgelet transform, we introduce a new name (i.e., MLDA) to emphasize the fundamental difference between the two approaches. Minimizing the mean-squared error, a wedgelet represents the most coherent regions in the corresponding dyadic square, that is, the regions with the minimal variance. On the other hand, an MLDA atom, supported by a dyadic square, represents not only the most coherent regions but also the regions with the maximum mean distance. One can easily imagine cases where minimizing the variances solely is insufficient for correct edge detection, especially for appearances in images that cannot be modeled with the Horizon model. In Fig. 1, we illustrate such a case where regions have almost identical means but different variances. Unlike the wedgelet transform designed for binary images, MLDA is capable of performing analysis of images represented in a multidimensional feature space (e.g., the RGB color space). In a more general setting of multidimensional data analysis, the MLDA extracts an "edge" between clusters of data, projecting the multidimensional data space onto the lower-dimensional MLDA representation.

MLDA exhibits the multiscale and localization properties of waveletes, but also offers additional information on alignments and spatial interrelationships among the extracted linear discriminants. In the following sections we explain these properties and demonstrate them with MLDA-based algorithms for analyzing geometrical structures in images.

Fig. 1. The wedgelet transform fails to detect the edge correctly: (left) original image, (center) wedgelet, (right) MLDA atom

2. MLDA FRAMEWORK

The MLDA framework consists of the following components:

1. **The MLDA atom** w is a piecewise constant function on either side of a linear discriminant that intersects the perimeter of a dyadic square S in vertices v_i and v_j. The discriminant (v_i, v_j) divides S into two regions R_0 and R_1, in which w takes values μ_0 and μ_1 equal to the mean vector of pixel values[1] in R_0 and R_1, respectively.

2. **The MLDA dictionary** is a collection of all possible MLDA atoms w at a finite range of locations, orientations and scales. MLDA is performed searching through the MLDA dictionary for atoms that best represent the analyzed image with respect to multiple criteria. The search keys are the location and scale of S and a discriminant (v_i, v_j). Computational complexity requires that the set of vertices $\{v_k\}$ on the perimeter of each S be finite (e.g., a number of pixels apart). In our implementation, the distance between two adjacent vertices is constant, which results in a different number of linear discriminants for squares at different scales.

3. **The MLDA tree** $\mathcal{T}$ consists of MLDA atoms generated in MLDA. Starting from S of size equal to the analyzed image, MLDA decomposes the image into children dyadic squares until optimization criteria are met. The MLDA tree $\mathcal{T}$ is incomplete, because generating atoms stops at different scales for different locations in the image. The leaf nodes of $\mathcal{T}$ store the final MLDA representation of an image, as illustrated in Fig. 2.

4. **The MLDA graph** is a graph plot of the MLDA representation of an image, where edges are linear discriminants and nodes are their corresponding vertices, as depicted in Fig. 2.

5. **MLDA-based algorithms** exploit information stored in the MLDA tree and graph. The tree structure gives rise to a host of multiscale analysis methods that account for μ_0 and μ_1 values of each w at various scales—the algorithms already introduced in the wavelet literature [5]. On the other hand, MLDA algorithms that use the MLDA graph go beyond the wavelet framework. These algorithms are capable of examining spacial interrelationships of linear discriminants, such as connectedness, collinearity and other properties of curves in an image.

We propose that the most important optimization criteria for MLDA are **discrimination** and **parsimony**. The pattern recognition literature abounds with various criterion functions for computing the best linear discriminant. In our implementation, we seek a direction (v_i, v_j), characterized by the maximum Mahalanobis distance between R_0 and R_1 in S, as

$$(v_i, v_j): \; d = \max_{(v_i, v_j)} \{(\mu_0 - \mu_1)^T (\Sigma_0 + \Sigma_1)^{-1} (\mu_0 - \mu_1)\}, \quad (1)$$

where Σ_0 and Σ_1 denote covariance matrices of R_0 and R_1, respectively. The computational cost of an exhaustive search over a finite set of linear discriminants $\{(v_i, v_j)\}$ can be reduced by updating the relevant statistics only with pixel values of delta regions (areas between two consecutive candidate linear discriminants).

As the size of S decreases, we achieve better piece-wise linear approximation of boundaries between regions in an image. Therefore, an analyzed image is decomposed into dyadic squares of varying sizes which results in the MLDA tree $\mathcal{T}$. To control the generation of children dyadic squares, we impose the next optimization criterion, parsimony, as a counter-balance to accuracy. We define a cost function to measure the parsimony of $\mathcal{T}$ as

$$R(\mathcal{T}) = \sum r(w_\ell) + \alpha |\mathcal{T}| \;, \quad (2)$$

where $r(w_\ell)$ is the inverse of the Mahalanobis distance computed for the corresponding leaf node w_ℓ, $r(w_\ell) = 1/d$, $|\mathcal{T}|$ denotes the number of terminal nodes in $\mathcal{T}$, and α represents the complexity cost per terminal node. Clearly, an exhaustive search in tree space for the minimum cost function is computationally prohibitive. Therefore, we implement the one-step optimal procedure analogous to the algorithm proposed in [6].

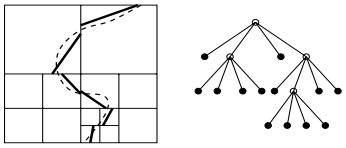

Fig. 2. (left) The MLDA graph: the dashed line depicts the actual curve; (right) the corresponding MLDA tree

Our experimental results suggest that no single stopping rule, which limits the size of $\mathcal{T}$, yields a satisfactory image representation. Therefore, instead of stopping at different terminal nodes, we continue MLDA image decomposition until all leaf squares are small in size, resulting in a large tree. Then, we selectively prune this large tree upward using the cost function $R(\mathcal{T})$. From expression (2), it follows that we can regulate the pruning process by increasing α to obtain a finite sequence of subtrees with progressively fewer leaf nodes. First, for each node $w \in \mathcal{T}$, we determine α_w for which the cost of a subtree $\mathcal{T}_w$ is higher than the cost of its root node w, as follows:

$$\begin{aligned} R(\mathcal{T}_w) \geq R(w) \quad &\Rightarrow \quad \sum r(w_\ell) + \alpha_w |\mathcal{T}_w| \geq r(w) + \alpha_w \cdot 1 \\ &\Rightarrow \quad \alpha_w = \frac{r(w) - \sum r(w_\ell)}{|\mathcal{T}_w| - 1} \;. \end{aligned} \quad (3)$$

Then, the whole subtree $\mathcal{T}_w$ under the node w with the minimum value of α_w is cut off. Repeatedly, we recalculate α_w values and trim off the weakest links until the actual number of leaf nodes is equal to or less than the desired number of leaf nodes.

This pruning is computationally fast and requires only a small fraction of the total tree construction time. Starting with a complete tree, the algorithm initially trims off large subtrees with many leaf nodes. As the tree becomes smaller, the procedure tends to cut off fewer nodes at a time. In Fig. 3, we illustrate the efficiency in image representation of the fully optimized MLDA tree, as compared to the unpruned MLDA tree. While there is almost no degradation in accuracy with complexity-cost pruning, we achieved significant reduction in the number of terminal MLDA atoms.

With additional optimization criteria, it is possible to meet various application specific requirements. For instance, in [4], the authors assume that curves in images appear smooth and connected, and therefore they impose a geometrical model as an optimization criterion for deriving the wedgelet image representation. Nevertheless, keeping in mind that the human visual system is capable of integrating scattered image constituents into coherent global structures [2], we let the MLDA representation provide all the richness

[1] In this paper we assume that the analyzed image is represented in multidimensional feature space, where a feature vector is assigned to each pixel.

Fig. 3. The MLDA representation: (left) original image, (center) no pruning 1024 leaf nodes, (right) with pruning 253 leaf nodes

of the underlying geometrical structures and assign the role of "integrating" (e.g. analysis of the smoothness and connectedness of curves) to MLDA-based algorithms.

3. ALGORITHMS AND RESULTS

MLDA is particularly suitable for extracting perceptually important features such as edges, singularities, and periodic patterns through a range of scales. Herewith, we focus only on the problem of extracting curves in an image, since a more thorough treatment is beyond the scope of this paper.

The vision-based horizon tracking for flight control and stability of micro air vehicles (MAV), as described in [5], is based on the assumption that the horizon is a straight line. Thus, the crucial part of the horizon detection algorithm is extracting candidate lines in an image—a task almost impossible to achieve using wavelets for noise-degraded images. In Fig. 4, we show two typical flight images, captured from an on-board camera of a MAV, where the video noise introduced edges and ripples that may mislead the wavelet-based edge detection. In this case, the MLDA framework offers an efficient solution without any need for image denoising. First, the MLDA image representation is found. Then, straight lines in the image are extracted analyzing exhaustively all possible directions, determined by a finite set of points $\mathcal{V}$ along the perimeter of the image. Each pair of points $(v_i, v_j) \in \mathcal{V}$ defines a direction with a slope a_{ij} along which MLDA atoms are examined. If slopes a of linear discriminants of the examined atoms satisfy

$$|a - a_{ij}| < \Delta \,, \tag{4}$$

where Δ denotes a slope step between two neighboring directions and can be computed as

$$\Delta = \min\{|a_{ij} - a_{kl}|/(1 - a_{ij}a_{kl})\} \;\;,\;\; k,l \in \{i, i \pm 1\} \,, \tag{5}$$

then that linear discriminant is extracted as a part of the analyzed direction (v_i, v_j). Finally, comparing μ_1 and μ_2 values of the extracted MLDA atoms with the prior statistical models of sky and ground [5], the list of horizon candidates is reduced to the horizon solution, as illustrated in Fig. 4.

In order to justify our claim that MLDA-based algorithms for edge extraction outperform wedgelet and wavelet-based approaches, especially in the presence of video noise, we carried out a comparative study of the three algorithms[2] for horizon detection on the same dataset comprising more than 1000 noisy MAV flight images. By visual inspection, we marked the "correct" position of the horizon for each image and then computed the percentage of erroneously classified pixels by the MLDA, wedgelet, and wavelet-based algorithms, which we present in Table 1. It seemed to us inappropriate to average the classification results, because it was only for a small number of test images that the error increased. Finally, it is note worthy that the MLDA-based horizon detection algorithm is capable of processing a video stream of flight images in real time.

[2]For more details on the wavelet-based horizon tracking, the interested reader is referred to [5].

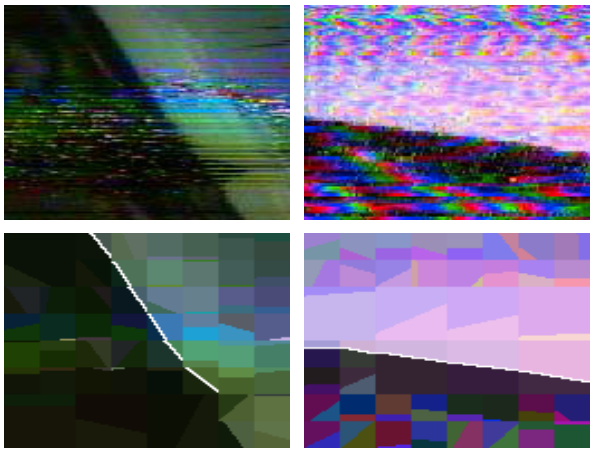

Fig. 4. Noise degraded MAV flight images

Table 1. Percentage of misclassified pixels for noisy flight images

MLDA	wedgelets	wavelets
5% - 17%	8% - 24%	12% - 31%

Further, we apply the described algorithm to solve the problem of detecting man-made structures in air images. Suppose, for example, that artificial structures appear as long straight edges, as the road in Fig. 5a. Thus, given this prior knowledge, the recognition problem can be solved implementing the MLDA-based edge detection. First, we exhaustively analyze all directions determined

(a) original image

(b) extraction of lines

(c) extraction of curves

Fig. 5. A noise degraded MAV flight image

by a finite set of points along the perimeter of the image, extracting linear discriminants. The directions most effectively covered with the extracted linear discriminants are candidate lines for the shapes of man-made objects. Criteria to choose the solution from the candidates can be based on examining various properties, such as: whether the lines are parallel or perpendicular. The two jagged white lines in Fig. 5b represent the two longest directions that are also most effectively covered with the extracted linear discriminants. It is most likely that these lines represent the road in the image.

Finally, we present an MLDA-based algorithm for extracting the minimum length curve between two fixed points in an image. Here, we implement the iterative-deepening A^* (IDA^*) search algorithm to optimize a path in the MLDA graph. It is well-known [7] that iterative-deepening search finds minimal cost paths with memory requirements that grow only linearly with the depth of the goal. Outward propagation from the starting point in the MLDA graph (i.e. the generation of new nodes in the IDA^* search tree) is controlled by a function

$$f(n) = g(n) + h(n) \,, \tag{6}$$

where $g(n)$ denotes the cost of the minimal cost path from the start node n_0 to the examined node n, and $h(n)$ denotes the actual cost of the minimal cost path between n and the goal node n_g. Recall that the exact positions in an image of all nodes (i.e. MLDA atoms) are known. Hence, we can estimate the true $h(n)$ with the following heuristic function:

$$\hat{h}(n) = d(n, n+1) + d_\ell(n+1) + d(n+1, n_g) \,, \tag{7}$$

where $n+1$ denotes a new node generated from the node n (i.e., a neighboring MLDA atom), then $d(n, n+1)$ stands for the minimum distance between the vertices of n and $n+1$, further, $d_\ell(n+1)$ denotes the length of the linear discriminant of $n+1$, and finally $d(n+1, n_g)$ stands for the minimum distance between vertices of the $n+1$ and n_g MLDA atoms. The controlling function $\hat{f}(n) = \hat{g}(n) + \hat{h}(n)$ is an accumulative sum of distances and, hence, a monotone nondecreasing function, assuring that the IDA^* search results in the minimum-length path.

The IDA^* algorithm executes a series of depth-first searches controlled by a cost cut-off value. We expand nodes in a depth-first fashion, backtracking whenever the $\hat{f}(n+1)$ value of a successor exceeds the cut-off. If this depth-first search does not terminate at the goal node n_g, then the cost cut-off value must be increased to start another depth-first search. The new cut-off value is set to the minimum $\hat{f}(n)$ value of the nodes visited but not expanded in the previous depth-first search.

Repetitive expansion of nodes in IDA^* hinders real-time implementation. On the other hand, the algorithm reduces memory requirements and improves implementation efficiencies of the depth-first search. In Fig. 5c, we illustrate the extraction of the minimum-length curve between two fixed points of interest. In this example, the start point is on the bottom left and the end point is on the upper right side of the image.

For space reasons, we left out substantial details as well as further examples of applied MLDA algorithms.

4. CONCLUSION

Currently well established image analysis approaches do not exhibit sufficient precision in detection and also do not provide sparse representation of image features. To deal with these problems, we proposed MLDA, a novel image analysis method, in this paper. We discussed the MLDA framework and showed that MLDA efficiently represents location, scale, orientation and elongation of image elements. We presented several MLDA-based algorithms for extracting curves in images as one application of the MLDA framework. Comparative study of the MLDA, wedgelet, and wavelet-based horizon detection in noisy images showed that MLDA outperformed the other two methods.

5. REFERENCES

[1] D.L.Donoho, "Wedgelets: Nearly-minimax estimation of edges," *Annals of Statistics*, vol. 27, no. 3, 1999.

[2] Spec. issue, "Natural Stimulus Statistics," *Network: Computation in Neural Systems*, vol. 12, 2001.

[3] D.J.Field, "Relations between the statistics of natural images and the response properties of cortical cells," *J. of Optical Soc. of America, A*, vol. 4, 1987.

[4] J.K.Romberg, M.Wakin, and R.G.Baraniuk, "Multiscale wedgelet image analysis: fast decompositions and modeling," in *IEEE ICIP 2002*, Rochester, NY, 2002.

[5] S.Todorovic and M.C.Nechyba, "Sky/ground modeling for autonomous MAV flight," in *IEEE ICRA 2003*, Taiwan, 2003.

[6] L.Breiman, J.Friedman, R.Olshen, and C.Stone, *Classification and Regression Trees*, Wadsworth, Belmont, CA, 1984.

[7] N. J. Nilsson, *Artificial Intelligence: a New Synthesis*, Morgan Kaufmann, San Francisco, CA, 1998.