

Image Multiresolution Discriminant Analysis for Vision-Guided Stability of Micro Air Vehicles

Sinisa Todorovic, Michael C. Nechyba, Antonio A. Arroyo

Dept. of Electrical and Computer Engineering,

University of Florida, Gainesville, FL 32611-6200

sinisha@mil.ufl.edu, nechyba@mil.ufl.edu, arroyo@mil.ufl.edu

Abstract—Recently, we have successfully implemented and tested a vision based horizon-tracking algorithm for flight stability and autonomy in Micro Air Vehicles (MAVs). Occasionally, this algorithm fails in scenarios where the underlying Gaussian assumption for the sky and ground appearances is not appropriate. To improve its performance, especially in the presence of video noise, we consider a novel image analysis tool—namely, multiresolution linear discriminant analysis (MLDA) that efficiently detects and economically represents edges in images. 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. With this approach, not only do we improve sky/ground segmentation results, but also enhance MAV’s potentials for surveillance and monitoring tasks.

I. INTRODUCTION

In this paper, we describe a novel image analysis tool that will allow us to segment reliably sky from ground in flight images and video. This goal was inspired by our previous work in horizon tracking for Micro Air Vehicles (MAVs) [1], [2]. In that work, we developed a real-time, vision-based horizon detection and tracking algorithm for MAVs equipped with on-board video cameras. With this system, we were able to achieve self-stabilized and autonomous flights of MAVs, without any additional inertial or rate sensors. We resorted to vision-based control, since such inertial and rate sensors typically do not yet have the requisite accuracy at the miniature scale required for MAVs, where weight of sensors and other components is of paramount importance.

Overall, the horizon tracking algorithm works well, especially when the sky and ground distributions are relatively coherent. Occasionally, however, horizon detection fails in scenarios where the underlying Gaussian assumption for the sky and ground appearances is not appropriate. Due to the noise degradation and/or complex scene appearances of sky and ground in images, this may be an incorrect assumption with potentially fatal consequences to the flight vehicle. Thus, given the goal to precisely detect a boundary between sky and ground regions (a requirement for MAV’s flight control and stability), careful consideration must be

given to selecting a reliable edge-extraction method.

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 [3], [4] 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 wavelets, also, characteristics that account for concepts beyond the wavelet framework.

Recently, the wedgelet transform has been proposed [5], 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 [5]–[7] 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 homogeneous regions in an image. Therefore, we propose a substantially different optimization criterion 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. 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 wavelets, 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.

II. 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 values¹ 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.

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

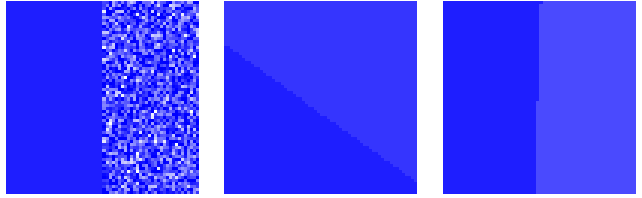


Fig. 1. The wedgelet transform fails to detect the edge correctly: (left) original image, (center) wedgelet, (right) MLDA atom

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 stop 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 naturally to a host of multiscale analysis methods which account for μ_0 and μ_1 values of each w at various scales. This type of algorithms has already been introduced in the wavelet literature (e.g., Hidden Markov Tree modeling for image segmentation [2]). There is but one noteworthy difference – namely, that the incomplete MLDA tree requires less processing, as opposed to a complete wavelet tree that includes a multitude of insignificant wavelet coefficients. The second type of algorithms goes beyond the wavelet framework analyzing spacial interrelationships of linear discriminants. Using the MLDA graph, these algorithms can efficiently examine 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) , char-

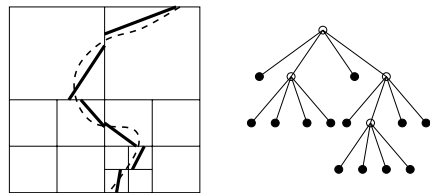


Fig. 2. (left) The MLDA graph: the dashed line depicts the actual curve; (right) the corresponding MLDA tree



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

acterized 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 piecewise 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 [8].

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:

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

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 [6], 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, we let the MLDA representation provide all the richness of the underlying geometrical structures and assign the role of “integrating” (e.g. analyzing the smoothness and connectedness of curves) to MLDA-based algorithms.

III. MLDA-BASED ALGORITHMS

As previously mentioned, MLDA gives rise to computer vision algorithms difficult or even impossible to implement within the wavelet framework. 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, with the emphasis on the MAV application, since a more thorough treatment is beyond the scope of this paper.

The existing vision-based horizon tracking for flight control and stability of our MAVs in flight is based on the assumption that the horizon is a straight line [1]. 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. 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

$$|a - a_{ij}| < \Delta, \quad (4)$$

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

$$\begin{aligned} \Delta &= \min\{|a_{ij} - a_{kl}|/(1 - a_{ij}a_{kl})\}, \\ k &\in \{i - 1, i, i + 1\}, \\ l &\in \{j - 1, j, j + 1\}, \end{aligned} \quad (5)$$

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 [2], the list of horizon candidates is reduced to the horizon solution, as illustrated in Fig. 4. The MLDA-based horizon detection algorithm is capable of processing a video stream of flight images in real time.

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

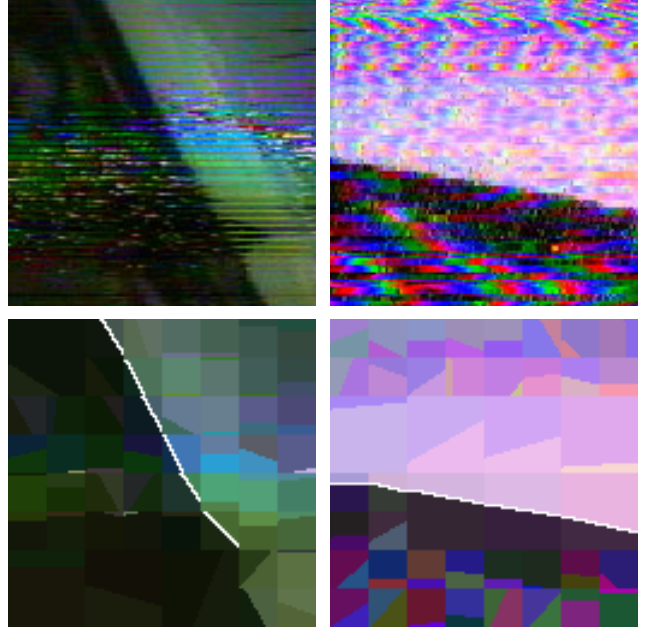


Fig. 4. Noise degraded MAV flight images

directions determined 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 [9], [10] 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

$$f(n) = g(n) + h(n), \quad (6)$$

where $g(n)$ denotes the cost of a minimal cost path from the start node n_0 to the current node n , and $h(n)$ denotes the actual cost of the minimal cost path between n and the goal node n_g . Clearly, $f(n_0) = h(n_0)$ for the start node. Recall that the exact positions in an image of all nodes (i.e. MLDA atoms) are known. Hence, we can estimate

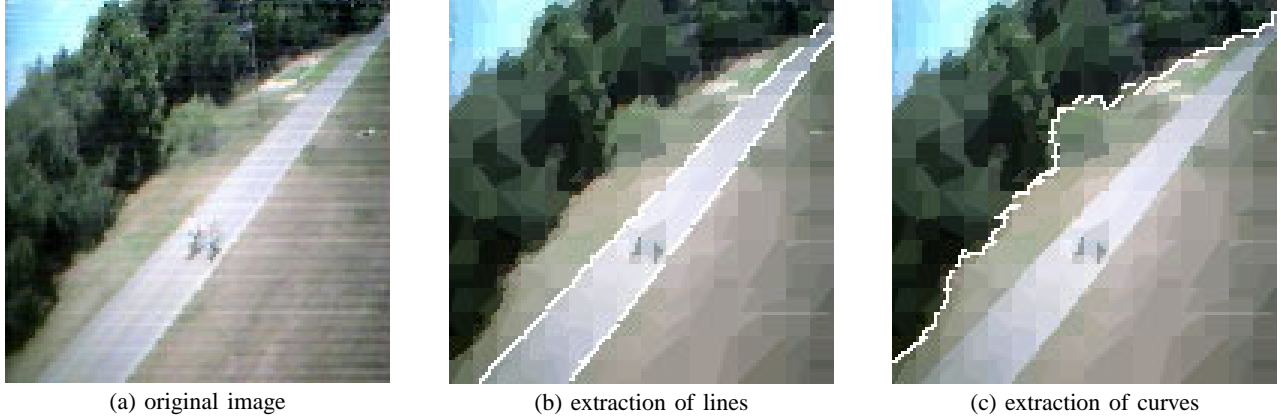


Fig. 5. Noise degraded MAV flight image

the true $h(n)$ with the following heuristic function:

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

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. In the first search, we establish a cost cut-off value equal to $\hat{f}(n_0) = \hat{g}(n_0) + \hat{h}(n_0) = \hat{h}(n_0)$. We expand nodes in a depth-first fashion, backtracking whenever the $\hat{f}(n+1)$ value of a successor exceeds the cut-off value. 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. The search ends when the MLDA atom n_g is reached.

The greatest drawback of IDA^* is its inherent repetitive expansions of nodes, which is the impediment for real time applications. On the other hand, there are fair trade-offs involving the reduced memory requirements and the 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 down left and the end point is on the up right side of the image.

IV. CONCLUSION

Considerations to improve the existing horizon-tracking algorithm for flight stability and control of Micro Air

Vehicles (MAVs) motivated our research of new image analysis tools. Currently well established methods 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. Then, to demonstrate its expressiveness, we presented MLDA-based algorithms for extracting curves in images, as one application of the MLDA framework. We showed that employing MLDA for the horizon-tracking yielded excellent sky/ground segmentation results in the presence of video noise. With MLDA, not only did we improve the horizon-tracking performance, but we also remained within the real-time constraints. Further, we presented a possible scenario, where a MAV was supposed to recognize a man-made object, namely a road. Also, we considered extraction of the minimum-length path in the MLDA graph between two points of interest. These examples illustrated some of the MAV's surveillance tasks, for which MLDA offered an efficient solution.

All the considered algorithms in this paper arise from the MLDA graph analysis. In our future research, we will focus on incorporating the multi-scale information, stored in the MLDA tree, into prospective computer vision algorithms which could lead to greater autonomy and versatility of our MAVs.

V. REFERENCES

- [1] S.M.Ettinger, M.C.Nechyba, P.G.Ifju, and M.Waszak, "Vision-guided flight stability and control for Micro Air Vehicles," in *IEEE IROS 2002, Proc.*, Lausanne, Switzerland, 2002.
- [2] S.Todorovic and M.C.Nechyba, "Adaptive feature selection for Hidden Markov Tree models of complex image classes," submitted to *IEEE Inter. Conf. CVPR 2003*, Madison, WI, 2003.
- [3] Spec. issue, "Natural Stimulus Statistics," *Network: Computation in Neural Systems*, vol. 12, 2001.

- [4] 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.
- [5] D.L.Donoho, "Wedgelets: Nearly-minimax estimation of edges," *Annals of Statistics*, vol. 27, no. 3, 1999.
- [6] J.K.Romberg, M.Wakin, and R.G.Baraniuk, "Multiscale wedgelet image analysis: fast decompositions and modeling," in *IEEE ICIP 2002, Proc.*, Rochester, NY, 2002.
- [7] R. L. Claypoole, Jr. and R. G. Baraniuk, "A multiresolution wedgelet transform for image processing," in *Proc. SPIE - The Int. Soc. for Optical Engineering*, San Diego, CA, Aug. 2000.
- [8] L.Breiman, J.Friedman, R.Olshen, and C.Stone, *Classification and Regression Trees*, Wadsworth Inc., Belmont, CA, 1984.
- [9] N. J. Nilsson, *Artificial Intelligence: a New Synthesis*, Morgan Kaufmann, San Francisco, CA, 1998.
- [10] C. Powley, C. Ferguson, and R. Korf, "Depth-first heuristic search on a SIMD machine," *Artificial Intelligence*, vol. 60, 1993.