Sky/Ground Modeling for Autonomous MAV Flight

Sinisa Todorovic[1], Michael C. Nechyba[1] and Peter Ifju[2]
sinisha@mil.ufl.edu, nechyba@mil.ufl.edu and pgi@aero.ufl.edu

[1]Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL 32611-6200
[2]Department of Mechanical and Aerospace Engineering, University of Florida, Gainesville, FL 32611-6250

***Abstract*— Recently, we have implemented a computer-vision based horizon-tracking algorithm for flight stability and autonomy in Micro Air Vehicles (MAVs) [2]. Occasionally, this algorithm fails in scenarios where the underlying Gaussian assumption for the sky and ground appearances is not appropriate. Therefore, in this paper, we present a general statisical image modeling framework, which we use to build prior models of the sky and ground; once trained, these models can be incorporated into our existing horizon-tracking aglorithm. Since the appearances of the sky and ground vary enormously, no single feature is sufficient for accurate modeling; as such, we rely both on color and texture as critical features in our modeling framework. Specifically, we choose hue and intensity for our color representation, and the complex wavelet transform (CWT) for our texture representation. We then use Hidden Markov Tree (HMT) models, which are particularly well suited for the CWT's inherent tree structure, as our underlying statistical models over our feature space. With this approach, we have achieved reliable and robust image segmentation of flight images from on-board our MAVs as well as on more difficult-to-classify sky/ground images; furthermore, we demonstrate the generality of our modeling framework through another segmentation task.**

I. Introduction

In this paper, we seek to build statistical appearance models that will allow us to segment sky from ground in images and flight 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. Moreover, the horizon detection algorithm is bootstrapped by assuming that initially the sky occupies the upper part of the image. For complex mission scenarios, this may be an incorrect assumption with potentially fatal consequences to the flight vehicle. For example, we are currently working on deploying MAVs on munitions for post-impact bomb damage assessment. In this case, the MAV would separate from the munition prior to impact, and an upright attitude with respect to the ground cannot be guaranteed. Correct identification of sky and ground, therefore, becomes imperative.

While modeling the appearance of sky and ground regions in images may seem intuitively easy, it is, in fact, a very challenging task. Depending on lighting, weather, landscape, etc., the appearance of the sky and ground can vary enormously. Given the complex variations in our two image classes (i.e. sky and ground), careful consideration must be given to selecting sufficiently discriminating features and a sufficiently expressive modeling framework. Having experimented with color and texture features separately, we conclude that only the feature set that includes both color and texture clues enables accurate statistical modeling for our application [3]. Previous experiments also suggest that it is important to represent both local as well as regional interdependencies in the feature space. As such, we resort to wavelet-based multi-resolution analysis in the form of the Complex Wavelet Transform (CWT).

Given our feature selection, we then choose the Hidden Markov Tree (HMT) model [4], [5] as our statistical model, since it is particularly well suited to the CWT's inherent tree structure. This choice of model imposes Markov dependencies on the states of both color values and wavelet coefficients at adjacent scales of the pyramidal multi-resolution structure. We train the HMTs with the EM algorithm [6] to obtain a small set of parameters that fully characterize the likelihoods of the two image classes at different scales. Finally, we fuse the posterior likelihoods at each scale, analogous to Choi's [7] interscale fusion approach, and perform Bayesian segmentation.

Our approach is distinguished from others, which use wavelets exclusively, by the inclusion of color in the HMT model structure. Incorporating color introduces a number of differences between our models and those in the literature. Moreover, the design of our statistical models was guided by real-time requirements of our MAV flight system, leading to certain design choices that may be sub-optimal if real-time processing constraints had not been an issue. Although it may appear that our vision algorithm is computationally

complex, we have come very close to meeting real-time requirements for our MAVs. Correct segmentation of a 128×128 resolution image takes only a fraction of a second on an Athlon 900MHz PC.

While this work was driven initially by our goal to model sky and ground appearances, our modeling framework does not restrict our work to this single case. To illustrate this, we present image segmentation results not only for sky/ground separation, but also for segmentation of *cut* and *uncut* grass regions; such segmentation capability may be useful, for example, for an autonomous lawnmower system.

Below, we give an overview of this paper. In Section II, we explain our choice of feature space, reviewing the most important aspects of the HSI color space and properties of the CWT. Next, in Section III, we describe the HMT model and Bayesian multiscale segmentation. Then, in Section IV, we present several examples of sky/ground and *cut*/*uncut* grass segmentation. Finally, we conclude with a discussion of our experimental results.

II. Feature space

For our statistical models, we seek to identify features that lead to improved segmentation performance without unnecessarily increasing computational complexity. As we have already mentioned, color or texture clues by themselves yield poor segmentation results [3]; therefore, below we consider a feature space that spans both color and texture domains.

A. Color

The color information in a video signal is usually encoded in the RGB color space. Unfortunately, the R, G and B color channels are highly correlated; therefore, we choose the HSI space as a more appropriate color representation for statistical modeling [8]. In order to simplify our feature space, we examine the Mahalanobis distances for the hue (H), saturation (S) and intensity (I) values in sky and ground training images. Denoting μ as the sample mean and σ^2 as the sample variance, we compute:

$$
\begin{aligned}
d_H^2 &= \frac{(\mu_H^{sky} - \mu_H^{ground})^2}{(\sigma_H^{sky})^2 + (\sigma_H^{ground})^2}, \\
d_S^2 &= \frac{(\mu_S^{sky} - \mu_S^{ground})^2}{(\sigma_S^{sky})^2 + (\sigma_S^{ground})^2}, \\
d_I^2 &= \frac{(\mu_I^{sky} - \mu_I^{ground})^2}{(\sigma_I^{sky})^2 + (\sigma_I^{ground})^2},
\end{aligned} \tag{1}
$$

and observe that for various training data sets d_H^2 and d_I^2 are consistently greater than d_S^2 [3]. Thus, to reduce computational complexity, we choose only the features H and I for our statistical model.

Next, we consider the representation of frequency, orientation and location of energy content in an image; in short, we want to define texture-based features. As such, we employ the wavelet transform, due it its inherent representation of texture at different scales and locations.

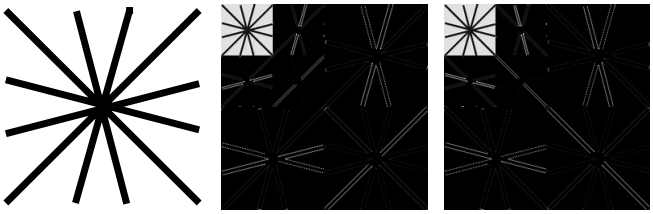

Fig. 1. The CWT is strongly oriented at angles $\pm15^\circ, \pm45^\circ, \pm75^\circ$. The original image (left) and the magnitude of the CWT coefficients: $15^\circ, 45^\circ, 75^\circ$ (center) and $105^\circ, 135^\circ, 165^\circ$ (right).

B. Complex Wavelet Transform

The 2-D Complex Wavelet Transform (CWT) essentially filters rows and columns of an image with a bank of complex bandpass filters, similar to the conventional Discrete Wavelet Transform (DWT) [9], [10]. Since, each coefficient contains a real and imaginary part, a redundancy of 2:1 is introduced for one-dimensional signals. For images, the redundancy increases to 4:1, since two adjacent quadrants of the spectrum are required to fully represent a real 2-D signal. This is achieved by additional filtering with complex conjugates of either row or column filters [11]–[13].

Despite its higher computational cost, we prefer the CWT over the DWT because of the CWT's following attractive properties. Kingsbury [11] has shown that the Dual-Tree CWT possesses near shift invariance, unlike the DWT, where small shifts in the input signal induces major changes in coefficient values. Also, the CWT's directional selectivity is greater, producing six bandpass subimages of complex coefficients at each level. The coefficients are strongly oriented at angles $\pm15^\circ, \pm45^\circ, \pm75^\circ$, as illustrated in Figure 1.

While it is known that the phase of CWT coefficients is less susceptible to noise corruption than the coefficient magnitudes [12], experimental results have shown that phase is not a good feature choice for sky and ground modeling [3]. Computing the phase of the CWT for orientation angles $\pm15^\circ, \pm45^\circ, \pm75^\circ$, yields virtually indiscernible subimages for sky and ground. Therefore, we consider only the magnitude of CWT coefficients in our representation of texture.

The magnitudes of CWT coefficients share the following properties of the DWT [4], [6], [9], [10]:

1) *Multi-resolution*: CWT represents an image at different scales of resolution in space.
2) *Clustering*: if the magnitude of a wavelet coefficient is large/small, then the magnitudes of the adjacent coefficients are very likely to also be large/small.
3) *Persistence*: large/small values of wavelet coefficients tend to propagate through scales.

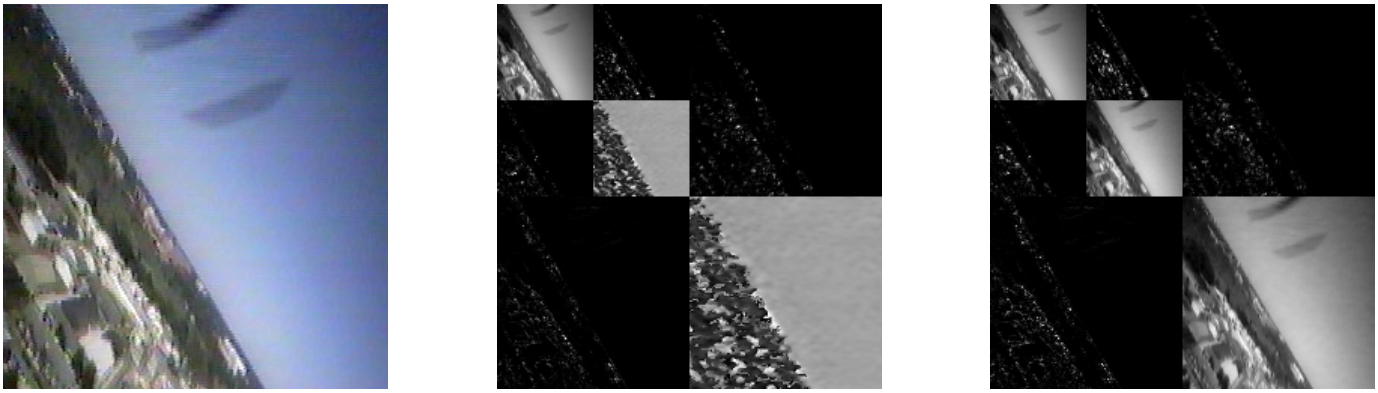

Fig. 2. The arrangement of the features: the original image (left), magnitudes of the 15° and 75° CWT and H values (center), magnitudes of the 105° and 165° CWT and I values (right).

These properties naturally give rise to the HMT statistical model, which helps us compute the distribution of pixels belonging to different image classes (as described in the next section).

To see which sets of orientations tend to be the most discriminating between sky and ground, we once again experiment with the Mahalanobis distances between sky and ground coefficient magnitudes belonging to subimages at different orientation. Computing $d^2_{15^\circ}$, $d^2_{-15^\circ}$, $d^2_{45^\circ}$, etc., similarly to the expressions in (1), we observe that for the available sky and ground training images $d^2_{45^\circ}$ and $d^2_{-45^\circ}$ are consistently the least significant. Therefore, our complete feature space is defined by the H and I color features and the subimages with orientation $\pm 15^\circ$ and $\pm 75^\circ$. To benefit from the multiscale presentation of the CWT, we replace the missing $\pm 45^\circ$ subimages with H and I images instead, as shown in Figure 2. The H and I values at coarser scales are computed as the mean of the corresponding four values at the next higher-resolution scale. Hence, the H and I features also exhibit the *clustering* and *persistence* properties to some extent.

Next we describe the HMT model as an appropriate statistical framework for modeling our chosen feature set.

III. Hidden Markov Tree Model

The Hidden Markov Tree (HMT) structure can model both the *clustering* and *persistence* properties of the CWT coefficient magnitudes. It consists of a tree structure $\mathcal{T}$ that assigns a node to each coefficient[1] and connects mutually dependent nodes. Thus, every parent node from coarser scales is vertically connected with their four[2] children at the finer scale, as depicted in Figure 3. For instance, it is obvious from the figure that

$$\mathcal{T}_{15^\circ} = \mathbf{W}^0_{15^\circ} \cup \mathbf{W}^1_{15^\circ} \cup \cdots \cup \mathbf{W}^{(L-1)}_{15^\circ} \qquad (2)$$

[1]Here, coefficient refers to the magnitude of CWT coefficients and/or the H and I color values.

[2]Throughout the paper the CWT is assumed to be dyadic.

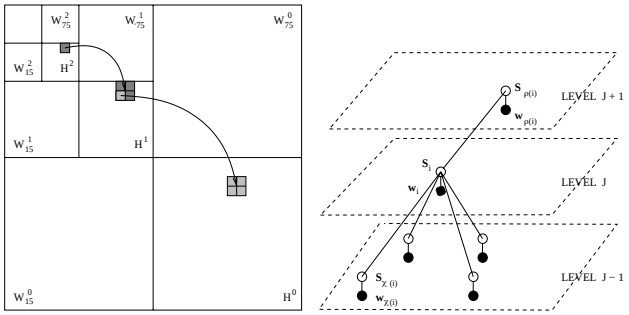

Fig. 3. The three-level CWT with H values: four adjacent coefficients at one scale have a unique parent belonging to the upper coarser scale. States $\mathbf{S}_i^J$ are depicted as white balls and coefficient values $\mathbf{w}_i^J$ as black balls.

Also, note from the figure that we assume that different features are mutually independent. In other words, connecting coefficients that belong only to the same feature, we obtain six mutually independent probability trees: $\mathcal{T}_{15^\circ}$, $\mathcal{T}_{75^\circ}$, $\mathcal{T}_{105^\circ}$, $\mathcal{T}_{165^\circ}$, $\mathcal{T}_H$, and $\mathcal{T}_I$.

It is worth noting that we tried to implement the Mixture Memory Markov Model, as proposed in [13], to account for the dependencies between features (i.e. probability trees). However, the slightly improved performance in image segmentation did not justify the substantial increase in processing time [3]. Also, we experimented with the HMT-2 model, developed in [14], where a coefficient depends on its two twin parents. Since the context-based fusion method used for Bayesian classification incorporates nine parents (not only two), segmentation performance did not improve with the HMT-2. Finally, we note that while we do not consider horizontal dependencies among nodes at the same scale, the *clustering* property is still well modeled, since adjacent coefficients at one scale have a unique parent.

In order to discuss HMT properties, we first need to introduce the following notation. A coefficient of a probability tree $\mathcal{T}_t$ at a scale J is denoted with $\mathbf{w}^J_{i,\mathcal{T}_t}$. A node *i* has one

parent node $\rho(i)$ and four children nodes $\chi(i)$, such that $J(\chi(i)) = J(i) - 1$ and $J(\rho(i)) = J(i) + 1$.

As is customary for HMTs [6], [15], we assign to each observable random variable (RV) $\mathbf{w}_{i,\mathcal{T}_t}^J$ a hidden RV, state $\mathbf{S}_{i,\mathcal{T}_t}^J$, which determines the marginal distribution of the observable coefficient value. The HMT imposes that $\mathbf{w}_{i,\mathcal{T}_t}^J$ is conditionally independent of all other RVs given its associated state $\mathbf{S}_{i,\mathcal{T}_t}^J$. Furthermore, $\mathbf{w}_{i,\mathcal{T}_t}^J$ is conditionally independent of the entire tree, given its parent state $\mathbf{S}_{\rho(i),\mathcal{T}_t}^{J+1}$. Note that the Markov structure is related to state RVs between scales and not to coefficient values.

If we assume an M-state Gaussian mixture density for the marginal distribution of $\mathbf{w}_{i,\mathcal{T}_t}^J$, the tree $\mathcal{T}_t$ is fully characterized by the following parameters:

1) The probability measure function of the root node: $P(\mathbf{S}_{0,\mathcal{T}_t}{=}m)$, $m{\in}[0, M-1]$.
2) The transition probability that $\mathbf{S}_{i,\mathcal{T}_t}^J$ is in a state *m* given that $\mathbf{S}_{\rho(i),\mathcal{T}_t}^{J+1}$ is in a state *n*: $a_{(J+1),J,\mathcal{T}_t}^{n,m}{=}P(\mathbf{S}_{i,\mathcal{T}_t}^J{=}m|\mathbf{S}_{\rho(i),\mathcal{T}_t}^{J+1}{=}n)$, $m{\in}[0, M{-}1]$, $n{\in}[0, M{-}1]$.
3) The mean and variance of $\mathbf{w}_{i,\mathcal{T}_t}^J$, given $\mathbf{S}_{i,\mathcal{T}_t}^J$ is in a state *m*: $\mu_{J,\mathcal{T}_t,m}$ and $\sigma_{J,\mathcal{T}_t,m}^2$, $m{\in}[0, M{-}1]$.

In order to simplify computations and to avoid the risk of overfitting the HMT model, we assume that the statistical parameters at the same scale are equal for all coefficients. Therefore, the model parameters are indexed by J , denoting that they are equal for all nodes *i* at the scale J. Finally, we group the parameters for all probability trees into a vector $\mathbf{\Theta}$.

Unlike in [6], [7], we do not assume zero mean values, since such an assumption would lead to substantial model error, especially for the H feature that takes on values in the interval $[0, 360]$. Also, much better image segmentation is obtained if the number of possible states M is greater than 2; since this introduces only a negligible increase in computation time, we let $M > 2$ (unlike in [6], [7]).

A. EM algorithm

Due to the Markov property and the assumption of probability tree independence, the most likely value for $p(\mathbf{w}_i^J|\mathbf{\Theta})$, can be computed by maximizing the joint likelihood:

$$\begin{aligned} p(\mathbf{w}_i^J, \mathbf{S}_i^J|\mathbf{\Theta}){=}&\prod_{\mathcal{T}_t} p(\mathbf{w}_{i,\mathcal{T}_t}^J, \mathbf{S}_{i,\mathcal{T}_t}^J|\mathbf{\Theta}) \ , \\ {=}&\prod_{\mathcal{T}_t} p(\mathbf{w}_{i,\mathcal{T}_t}^J|\mathbf{S}_{i,\mathcal{T}_t}^J, \mathbf{\Theta}) P(\mathbf{S}_{i,\mathcal{T}_t}^J|\mathbf{\Theta}) \ , \\ {=}&\prod_{\mathcal{T}_t} p(\mathbf{w}_{i,\mathcal{T}_t}^J|\mathbf{S}_{i,\mathcal{T}_t}^J, \mathbf{\Theta}) \cdot \\ &\cdot P(\mathbf{S}_{0,\mathcal{T}_t}{=}m|\mathbf{\Theta}) \prod_J a_{J+1,J,\mathcal{T}_t}^{n,m} \quad , \end{aligned} \tag{3}$$

where *t* takes on the values $\{\pm 15^\circ, \pm 75^\circ, H, I\}$.

The last expression shows all the HMT parameters which must be learned from observations. For training the HMT model, we implement the iterative Expectation-Maximization (EM) algorithm, as proposed in [6]. In the *E* step, the state information is propagated throughout the tree by means of the *upward-downward* algorithm. Here, at step *l* of the algorithm, the expectation value of the log-likelihood from (3) is computed as follows:

$$\begin{aligned} Q(\mathbf{\Theta}, \mathbf{\Theta}^l){=}&E_{\mathbf{S}_i^J}\left[ln\ p(\mathbf{w}_i^J, \mathbf{S}_i^J|\mathbf{\Theta}) \mid \mathbf{w}_i^J, \mathbf{\Theta}^l \right] \ , \\ {=}&\sum_m ln\ p(\mathbf{w}_i^J, \mathbf{S}_i^J|\mathbf{\Theta}) P\left(\mathbf{S}_i^J{=}m|\mathbf{w}_i^J, \mathbf{\Theta}^l\right) . \end{aligned} \tag{4}$$

Then, in the *M* step, we compute

$$\mathbf{\Theta}^{l+1} = arg \max_{\mathbf{\Theta}} Q\left(\mathbf{\Theta}, \mathbf{\Theta}^l\right) \ . \tag{5}$$

It has been proved that increasing the *Q*-function is sufficient to increase the likelihood $p(\mathbf{w}_i^J|\mathbf{\Theta})$ [16]. We are not concerned with the convergence rate in the training process, because our data base contains long sequences of similar sky and ground images. Hence, $\mathbf{\Theta}^l$, computed for one image, is used as the input to compute $\mathbf{\Theta}^{l+1}$ for the next image of the training data base. Finally, after processing all sky training images, we obtain $\mathbf{\Theta}^{sky}$, and similarly, for ground, $\mathbf{\Theta}^{ground}$.

Thus, the EM algorithm gives us the likelihoods of all coefficients at all scales for a given class, say sky, as follows:

$$p(\mathbf{w}_i^J|\mathbf{\Theta}^{sky}) = \sum_m p(\mathbf{w}_i^J, \mathbf{S}_i^J{=}m|\mathbf{\Theta}^{sky}). \tag{6}$$

Consequently, we are able to perform Bayesian classification at all scales, without significant computational overhead.

B. Multiscale Bayesian Segmentation

Most segmentation algorithms employ a classification window of some size, which provides statistical information to a classifier. A large classification window produces accurate segmentation of large, homogeneous regions, but poor results along their boundaries. On the other hand, a small window yields unreliable classification. In our case, we require not only recognition of the sky and ground regions, but als the detection of the horizon with as much accuracy as possible. Therefore, both large and small scale neighborhoods should be analyzed. Naturally, to benefit from our already trained HMT model, we again resort to its multiscale structure to perform segmentation. Thus, we implement a multiscale segmentation algorithm, similar to the one developed in [7]; however, we do not employ contexts, and thus achieve much faster processing times. Also, the dyadic squares of [7], in our case diminish in size to only one coefficient. Both of these differences in approaches were guided by our requirement for precise horizon-line detection and real-time processing constraints.

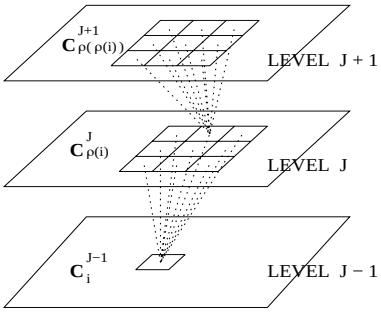

Fig. 4. The Markov tree structure: a class label $\mathbf{C}_i^J$ at the scale J has nine parents $\mathbf{C}_{\rho(i)}^{J+1}$ at the coarser scale $J+1$.

First, for $\mathbf{\Theta}^{sky}$ and $\mathbf{\Theta}^{ground}$, we compute likelihoods, as given by (6), for all coefficients at all scales. Then, by the maximum-likelihood (ML) criterion, we assign to each coefficient the label *sky* or *ground*. Now we need to combine the ML results at all scales for the final segmentation. For that, we form a new Markov tree structure, but this time in reverse order, as shown in Figure 4.

Here, a class label, denoted with $\mathbf{C}_i^J$, plays a role analogous to the hidden state $\mathbf{S}_i^J$ in the HMT. In the tree[3], every class label $\mathbf{C}_i^J$ has nine corresponding parents at the coarser scale $\mathbf{C}_{\rho(i)}^{J+1}$. To facilitate computation, we assume that $\mathbf{C}_i^J$ class labels (i.e. ML results) are mutually independent. Our Markov tree is completely characterized with the *a priori* probabilities of the class labels at the coarsest scale, and also with the transition probabilities between parent and children labels. These values must be learned from the given training images. To do this, we once again implement the EM algorithm to estimate the prior and transition probabilities. Finally, using the EM results for sky and ground images and employing the Markov chain rule, we obtain the probability measure function $P(\mathbf{C}_i^0 = \textit{class})$, $\textit{class} \in \{\textit{sky}, \textit{ground}\}$, for all coefficients at the finest 0-th level. These values are then used in a Bayes classifier to obtain the desired image segmentation.

IV. Results

For training the HMT model, we recorded two sets of 500 sky and ground images. We carefully chose the training sets to account for great variability within classes. After experimenting with different image resolutions, we found that the best trade off between processing time and performance was achieved for 128×128 resolution images. At that resolution, the training time on an Athlon 900MHz PC was less than 3 minutes for both classes. In Figures 5-7 we present segmentation results for three diverse sky/ground images; these results incorporated the Q-shift Dual-Tree CWT introduced in [17].

[3]Note that in this case there is only one Markov tree, whereas for the HMT there are six trees for six features.

Finally, we illustrate the generality of our algorithm in segmentation of *cut*/*uncut* grass regions (see Figure 8). These images[4] show a grass lawn from the perspective of a camera mounted on an autonomous lawn mower. Even at image resolutions as low as 64×64, we achieve satisfactory results at very fast processing speeds.

V. Conclusion

Segmentation of complex image classes, such as sky and ground, demand an elaborate consideration of class properties. Clearly, in some cases, color provides sufficient information for sky and ground detection. However, due to video noise and/or unfavorable class patterns, both color and texture clues are necessary for successful recognition.

In this paper, we first presented our choice of features, consisting of H and I values from the HSI color space, and CWT coefficients. Then, we showed the implementation of the HMT model and the training steps for obtaining its parameters. We further described how the learned paremeter set could be used for computing likelihoods of all nodes at all scales of our HMT. We then developed multiscale Bayesian classification for our application. We incorporated in our design results from the available literature, modifying the original algorithms for our purposes where appropriate. Most importantly, we incorporated color features into the HMT framework and designed the consequent classifier with real-time constraints in mind. Finally, we show sample classification results on diverse sky/ground images and on *cut*/*uncut* grass images.

References

[1] S. M. Ettinger, "Design and implementation of autonomous vision-guided Micro Air Vehicles," Master's thesis, UF, Gainesville, FL, 2001.

[2] S. M. Ettinger, M. C. Nechyba, P. G. Ifju, and M. Waszak, "Vision-guided flight stability and control for Micro Air Vehicles," in *Proc. IEEE Int. Conference on Intelligent Robots and Systems*, October 2002.

[3] S. Todorovic, "Image modeling using the tree-structured Hidden Markov Model," Master's thesis, UF, Gainesville, FL, 2002.

[4] J. K. Romberg, H. Choi, R. G. Baraniuk, and N. Kingsbury. (2002, Apr.) Hidden Markov Tree Models for Complex Wavelet Transforms. Rice DSP Publications. [Online]. Available: http://www-dsp.rice.edu/~jrom/publications.html

[5] H. Choi, J. Romberg, R. Baraniuk, and N. Kingsbury, "Hidden Markov Tree modeling of Complex Wavelet Transforms," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, Instanbul, Turkey, June 2000.

[6] M. S. Crouse, R. D. Nowak, and R. G. Baraniuk, "Wavelet-based statistical signal processing using Hidden Markov Models," *IEEE Transactions on Signal Processing*, vol. 46, no. 4, pp. 886–902, April 1998.

[7] H. Choi and R. G. Baraniuk, "Multiscale image segmentation using wavelet-domain Hidden Markov Models," *IEEE Transactions on Image Processing*, vol. 10, no. 9, pp. 1309–1321, Sept. 2001.

[8] H. D. C. et al., "Color image segmentation: advances and prospects," *Pattern Recognition*, vol. 34, pp. 2259–2281, 2001.

[9] S. Mallat, *A Wavelet Tour of Signal Processing*, 2nd ed. Academic Press, 2001.

[4]We thank Rand Chandler for giving us access to his image database.

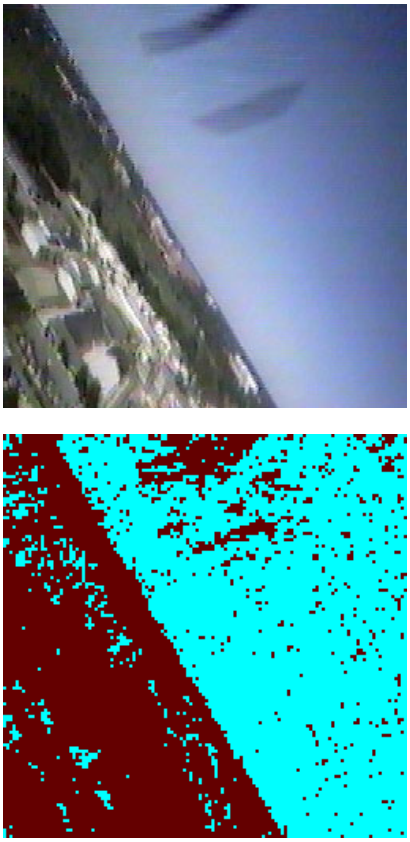

Fig. 5. A MAV's flight. The MAV's propeller blades are recognized as *ground*.

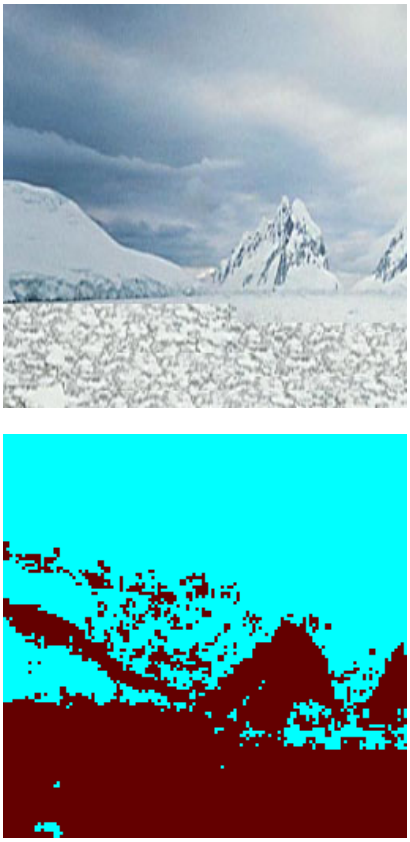

Fig. 6. Water surface with ice patches similar to clouds.

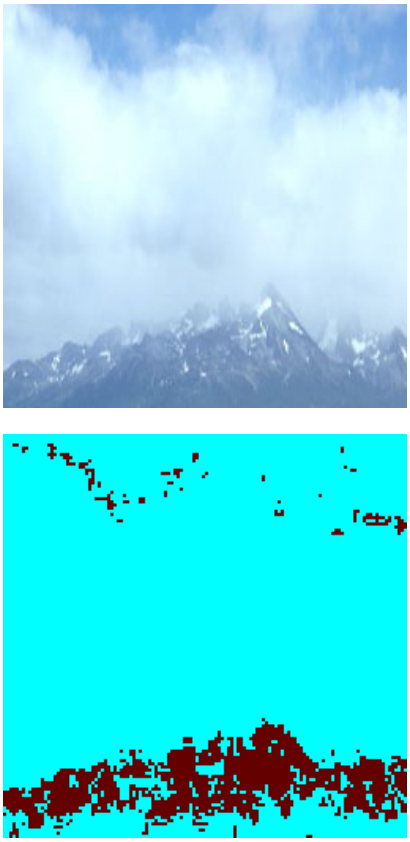

Fig. 7. A mountain view with predominant blue color and fuzzy horizon line.

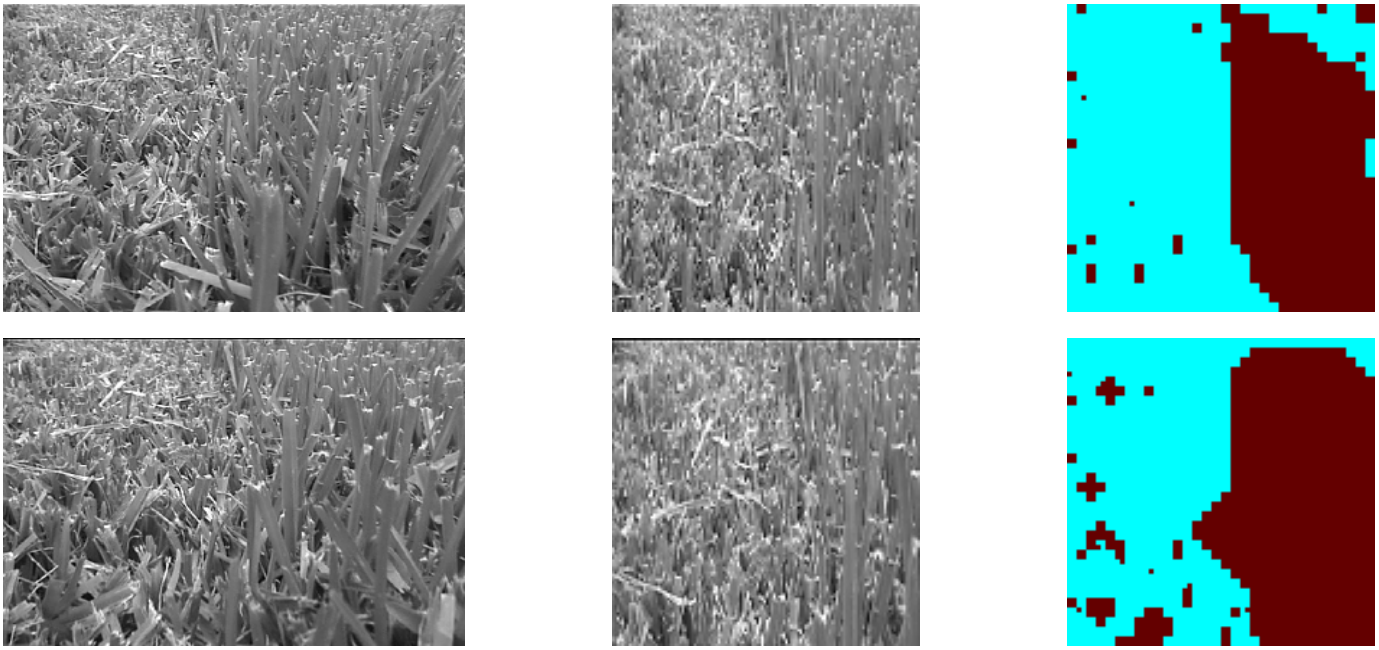

Fig. 8. Segmentation of *cut* and *uncut* grass regions. The original image (left), subsampled image (center), segmentation of the subsampled image (right).

[10] S. G. Mallat, “A theory for multiresolution signal decomposition: the wavelet representation,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 7, pp. 674–693, July 1989.

[11] N. Kingsbury, “Image processing with complex wavelets,” *Philosophical Transactions Royal Society London*, vol. 357, no. 1760, pp. 2543–2560, 1999.

[12] J.-M. Lina, D. Clonda, and B. Goulard. Statistical modeling with complex dyadic wavelets. Personal Web Page. [Online]. Available: http://www.crm.umontreal.ca/˜clonda/

[13] D. Clonda, J.-M. Lina, and B. Goulard, “Mixed memory model for image processing and modeling with complex Daubechies wavelets,” in *Proc. SPIE’s International Symposium on Optical Science and Technology*, San Diego, CA, Aug. 2000.

[14] G. Fan and X.-G. Xia, “Improved Hidden Markov Models in the Wavelet-Domain,” *IEEE Transactions on Signal Processing*, vol. 49, no. 1, pp. 115–120, Jan. 2001.

[15] L. R. Rabiner, “A tutorial on Hidden Markov Models and selected applications in speech recognition,” *Proceedings of the IEEE*, vol. 77, no. 2, pp. 257–286, Feb. 1989.

[16] G. J. McLachlan and K. T. Thriyambakam, *The EM algorithm and extensions*. John Wiley & Sons, 1996.

[17] N. Kingsbury, “Complex wavelets for shift invariant analysis and filtering of signals,” *Journal of Applied and Computational Harmonic Analysis*, vol. 10, no. 3, pp. 234–253, May 2001.