On the Use of Marginal Statistics of Subband Images

Joshua Gluckman*
Dept. of Computer Science
Polytechnic University
Brooklyn, NY 11201
jgluckma@poly.edu

Abstract

A commonly used representation of a visual pattern is the set of marginal probability distributions of the output of a bank of filters (Gaussian, Laplacian, Gabor etc...). This representation has been used effectively for a variety of vision tasks including texture classification, texture synthesis, object detection and image retrieval. This paper examines the ability of this representation to discriminate between an arbitrary pair of visual stimuli. Examples of patterns are derived that provably possess the same marginal statistical properties, yet are "visually distinct." These results suggest the need for either employing a large and diverse filter bank or incorporating joint statistics in order to represent a large class of visual patterns.

1. Introduction

The ability of a visual system to discriminate among a multitude of stimuli ultimately depends on the underlying representation of a visual pattern. Computing a large number of statistical measures from a set of filtered images is one commonly used representation. Many successful methods for object recognition, object detection, image retrieval, texture synthesis, and texture recognition have been developed based on such a representation [1] [2] [3] [4]. In this class of methods, a set of subband images is created through convolution with a bank of filters, Gaussian, Laplacian, Gabor, etc..., then statistical measures are computed from the subband images. A variety of statistical measures have been proposed including: parametric models, moments, entropies, histograms, and joint distributions. Detection and recognition is performed by classifying the statistical representation of a novel image while texture synthesis is performed by randomly sampling from the ensemble of images that match the statistics of a particular pattern. In order to understand the ability of these methods to scale to a large number of visual patterns it is important to examine the ability of the underlying representation to discriminate between any pair of visual stimuli.

In this paper we address the use of marginal statistics (histograms) of images filtered with Gaussian, Laplacian, derivative and Gabor functions. In particular, we derive examples of visual patterns that are "distinct" yet *provably* map to the same marginal probability density function of a large filter bank. The fact that these patterns exist does not imply that marginal distributions are not important or useful. To the contrary, marginal statistics have been used to successfully synthesize and classify a variety of textures and patterns. However, the results in this paper do suggest the need for certain classes of filters, a large filter bank and/or the use of joint statistical information.

One of the first models for representing a visual pattern was proposed by Julesz, who suggested that co-occurence statistics of k-tuples of pixels might explain texture perception [5]. The model was later disproved by Julesz and others by constructing patterns that were distinct to a human observer but shared the same high order co-occurence statistics [6]. Later, Markov random fields were proposed [7]. However, such models were difficult to deal with due to the high computational complexity. More recently texture recognition and synthesis has focused on filtering theory. Faugeras and Pratt suggested representing textures by the marginal statistics of images after applying a bank of filters. This line of research has been motivated by both neurophysiological evidence that the human visual system decomposes the retinal image into subband channels and psychophysical evidence that textures sharing similar marginal distributions are difficult to discriminate [8] [9] [10]. Heeger and Bergen developed an algorithm for synthesizing textures that match the histograms of a target texture across the levels of a Steerable pyramid (oriented filters) [11]. Zhu et al. advocated using the marginal distributions (histograms) of a large filter set and demonstrated synthesis of a variety of visual patterns and textures [12]. An important theoretical result linking these filtering methods to Markov random field models was shown by Zhu et al. [13]. Work by De Bonet and Viola has shown texture synthesis using the joint statistics of filter outputs [14]. In addition, Portilla and Simoncelli have also expressed the need for joint statistics when synthesizing textures [15]. Recently, good results on texture classification have been obtained using "textons", which are clusters in the joint statistical space of filter outputs [16] [17] [18].

*This work was supported by a NSF ITR award no. IIS-0219078

Proceedings of the Ninth IEEE International Conference on Computer Vision (ICCV 2003) 2-Volume Set
0-7695-1950-4/03 $17.00 © 2003 IEEE

IEEE COMPUTER SOCIETY

While it is clear that statistical measures of subband images are a useful representation of visual patterns, questions regarding the expressive power of any particular statistic and any particular set of filters need to be answered in order to design efficient and robust visual systems.

Previously, Zhu et al. have shown that the histograms of filtered images are sufficient to uniquely characterize an image, however the proof requires an uncountably infinite number of filters [12]. Hadjidemetriou et al. derived the class of continuous transformations that preserve the histogram of an unfiltered image [19]. They also studied the histograms of Gaussian filtered images and were able to demonstrate sensitivity to many texture properties [20].

The main contribution of this paper is a study of the ability of marginal statistics of commonly used filters to discriminate visual patterns. For several classes of filters we show the existence of patterns that map to the same marginal distributions but are "visually distinct." Furthermore, they are not rigid or even affine transformations of each other. Along the way, several novel results are shown relating the marginal statistics to the image frequencies.

2. Marginal Statistics

Given an image $f(x, y)$ defined on the continous image domain $\Omega \subset \Re^2$ and a set of filters $g = \{g_1, \ldots, g_k\}$ the subband images $\{f_1, \ldots, f_k\}$ are formed by the convolution $f_\alpha = g_\alpha * f$. The marginal probability distributions of the filter outputs consider each subband image in isolation and are defined by

$$p_\alpha(i; f) = \frac{1}{|\Omega|} \int_\Omega \delta(f_\alpha - i) d\Omega, \tag{1}$$

where δ is the Dirac delta function. For a fixed image size the marginal distribution is equivalent to the histogram. Joint distributions are functions of the output of several or all filters and given by

$$p(i_1, i_2, \ldots, i_k; f) = \frac{1}{|\Omega|} \int_\Omega \prod_{\alpha=1}^{k} \delta(f_\alpha - i_\alpha) d\Omega. \tag{2}$$

For a fixed set of filters the joint distribution contains more information than the marginal distributions. However, the storage and comparison costs of joint distributions (multi-dimensional histograms) grow exponentially. Clearly, we would like to fully understand the expressive power of marginal distributions before resorting to joint statistics.

The question addressed in this paper is do patterns f and f' exist that have the same marginal distributions for a large, perhaps infinite set of filters g, meaning $p_\alpha(f) = p_\alpha(f'), \forall \alpha : g_\alpha \in g$. Furthermore, we place the constraint that f and f' are not related by an affine transformation. It is important to note that matching the marginal distributions implies that any statistical measure derived from the marginal distributions, such as entropies, range, median, will also be identical. To answer this question we consider the class of functions that are the sum of cosine functions occurring at integer frequencies with an arbitrary but fixed set of weights.

Definition Given a set of n weights $w = (w_1, \ldots, w_n) : w_i \in \mathbb{R}$, define $F(w)$ to be the set of functions such that for all $u = (u_1, \ldots, u_n) \in \mathbb{Z}^n$ and $v = (v_1, \ldots, v_n) \in \mathbb{Z}^n$

$$\sum_{i=1}^{n} w_i \cos(u_i x + v_i y) \in F(w).$$

In the following section we derive sufficient conditions to ensure a pair of functions $f, f' \in F(w)$ have the same or similar histograms, $p(f) = p(f')$. The conditions are independent of the weights w and only depend on the choice of integer frequency pairs $\{(u_1, v_1), \ldots, (u_n, v_n)\}$ for f and $\{(u'_1, v'_1), \ldots, (u'_n, v'_n)\}$ for f'. Then, this result is extended by deriving additional constraints that if satisfied ensure f and f' have the same histograms after convolution with a bank of filters.

3. Moments

The density function (1) is not defined directly on the image f therefore we work with the moments of the distribution. Let M^k be the k^{th} moment of the distribution (1),

$$M_k(f) = \int i^k p(i) di = \frac{1}{|\Omega|} \int_\Omega f(x, y)^k d\Omega.$$

Provided the existence of the moment generating function the moments uniquely represent the probability distribution [21]. For images, only a finite number of moments are needed to represent the distribution due to the presence of limited dynamic range, quantization and noise. Thus, sufficient conditions are derived for ensuring the first k moments are identical for a pair of functions f and f' in $F(w)$ implying that $p(f) \approx p(f')$.

Dropping the normalization constant, the moments for a function in $F(w)$ are

$$M_k = \int_\Omega \left(\sum_{i=1}^{n} w_i \cos(z_i) \right)^k d\Omega, \tag{3}$$

where $z_i = u_i x + v_i y$. Applying the multinomial theorem[1]

$$M_k = \int_\Omega \sum_{k_1, k_2, \ldots, k_n} \frac{k!}{k_1! k_2! \cdots k_n!} \prod_{i=1}^{n} w_i^{k_i} \cos^{k_i}(z_i) d\Omega,$$

where the sum is taken over all of the non-negative integers such that $\sum_i^n k_i = k$. To simplify the notation this is written as

$$M_k = \int_\Omega \sum_{\kappa \in K} \mathcal{F}(\kappa, w) \prod_{i=1}^{n} \cos^{k_i}(z_i) d\Omega,$$

[1] A generalization of the binomial theorem

Proceedings of the Ninth IEEE International Conference on Computer Vision (ICCV 2003) 2-Volume Set
0-7695-1950-4/03 $17.00 © 2003 IEEE

with $\kappa = (k_1, \ldots, k_n)$ and $K = \{\kappa : \forall i k_i \geq 0, \sum_i^n k_i = k\}$. In order to remove the exponentiation the following substitution

$$\cos^a(z) = \sum_{i=0}^{a} C(i,a)\cos(iz),\ a \in \mathbb{Z}$$

is applied. We are not concerned with the actual values of $C(i,a)$ except to note that they are independent of z and have the following property

$$C(i,a) \neq 0 \Leftrightarrow \text{parity}(i) = \text{parity}(a). \tag{4}$$

Applying the substitution

$$M_k = \int_\Omega \sum_{\kappa \in K} \mathcal{F}(\kappa, w) \prod_{i=1}^{n} \left(\sum_{r_i=0}^{k_i} C(r_i, k_i)\cos(r_i z_i) \right) d\Omega.$$

Interchanging the product and the summation results in

$$M_k = \int_\Omega \sum_{\kappa \in K} \mathcal{F}(\kappa, w) \sum_{r \in R_\kappa} \prod_{i=1}^{n} C(r_i, k_i)\cos(r_i z_i) d\Omega,$$

where $r = \{r_1, \ldots, r_n\}$ and the summation is taken over the set $R_\kappa = \{r : 0 \leq r_i \leq k_i,\ \forall i\}$. Again simplifying the notation and moving the integral inside the summation

$$M_k = \sum_{\kappa \in K} \mathcal{F}(\kappa, w) \sum_{r \in R_\kappa} \mathcal{C}(r, \kappa) \int_\Omega \prod_{i=1}^{n} \cos(r_i z_i) d\Omega,$$

where $\mathcal{C}(r,\kappa) = \prod_{i=1}^{n} C(r_i, k_i)$. From (4) $\mathcal{C}(r,\kappa)$ has the following property

$$\mathcal{C}(r,\kappa) \neq 0 \Leftrightarrow \forall i\ \text{parity}(r_i) = \text{parity}(k_i). \tag{5}$$

The product is removed by repeatedly applying by the following trigonometric substitution

$$\cos(a)\cos(b) = \frac{1}{2}(\cos(a+b) + \cos(a-b)),$$

resulting in

$$M_k = \sum_{\kappa \in K} \mathcal{F}(\kappa, w) \sum_{r \in R_\kappa} \mathcal{C}(r,\kappa) \int_\Omega \frac{1}{2^{n-1}} \sum_{d_2, \ldots, d_n} \cos\left(\sum_{i=1}^{n} d_i r_i z_i \right) d\Omega,$$

where $d_1 = 1$ and $d_i = \pm 1$ for $i \geq 2$. The left summation inside the integral is taken over the 2^{n-1} possible assignments of the d_i. Note that each integral of the form

$$\int_\Omega \cos(r_1 z_1 \pm r_2 z_2 \pm \cdots \pm r_n z_n) d\Omega.$$

Recall that $z_i = u_i x + v_i y$ and $r_i, u_i, v_i \in \mathbb{Z}$ therefore if the integration is performed over $(-\pi, \pi)$ in both the x and y direction then each integral evaluates to 0 unless all of the terms sum to 0 in which case the integral evaluates to $|\Omega| = 4\pi^2$ thus,

$$\int_\Omega \cos\left(\sum_{i=1}^{n} d_i r_i z_i \right) d\Omega = \begin{cases} 4\pi^2 & \text{if } \sum_i^n d_i r_i z_i = 0 \\ 0 & \text{otherwise} \end{cases}$$

From this result the integral is removed and

$$M_k = \sum_{\kappa \in K} \mathcal{F}(\kappa, w) \sum_{r \in R_\kappa} \mathcal{C}(r,\kappa) \frac{1}{2^{n-1}} t(u, v, r). \tag{6}$$

Note that the $4\pi^2$ is removed because it cancels with the normalization constant that was dropped in (3). The function $t(u,v,r)$ is the number of solutions to the 2^{n-1} equations of the form

$$r_1 z_1 \pm r_2 z_2 \pm \cdots \pm r_n z_n = 0,$$

which is an integer in the range $(0, 2^{n-1})$. Note that t is a function of u, v, r because a particular equation,

$$d_1 r_1 z_1 + \cdots + d_n r_n z_n = 0,$$

is satisfied if and only if

$$\sum_{i=1}^{n} d_i r_i u_i = 0 \wedge \sum_{i=1}^{n} d_i r_i v_i = 0. \tag{7}$$

Equation (6) separates out the dependence of the moment function on the integer frequencies u and v and will be used to prove several results linking the marginal distributions to the image frequencies. From (7) the function $t(u,v,r)$ is determined by the solutions to a pair of Diophantine equations[2].

Definition Given a function f with the integer frequencies $(u_1, \ldots, u_n)$ and $(v_1, \ldots, v_n)$ define $A(f) = \{(m_1, \ldots, m_n)\} \subseteq \mathbb{Z}^n$ to be the set of solutions to the Diophantine equations

$$\begin{aligned} u_1 m_1 + u_2 m_2 + \cdots + u_n m_n &= 0 \\ v_1 m_1 + v_2 m_2 + \cdots + v_n m_n &= 0. \end{aligned} \tag{8}$$

Now, we are ready to show the conditions under which a pair of functions f and f' have the same moments.

Theorem 3.1 *Let f and f' be functions in $F(w)$. If $\forall c \in A(f) \ominus A(f')$ there exists an integer $\hat{k}$ such that $\sum_i^n |c_i| > \hat{k}$ then $\forall k \leq \hat{k}$, $M_k(f) = M_k(f')$.*

Proof We will prove the contrapositive statement. Thus, we assume there exists a $k \leq \hat{k}$ such that $M_k(f) \neq M_k(f')$. From (6) there must exist a $\kappa = (k_1, \ldots, k_n)$ and an $r = (r_1, \ldots, r_n)$ such that $t(u,v,r) \neq t(u',v',r)$ where

[2] An equation in which only integer solutions are allowed.

Proceedings of the Ninth IEEE International Conference on Computer Vision (ICCV 2003) 2-Volume Set
0-7695-1950-4/03 $17.00 © 2003 IEEE

$\sum_{i=1}^{n} k_i = k$ and $r_i \leq k_i$. Without loss of generality assume that $t(u, v, r) > t(u', v', r)$, which implies there must exist a $(d_1, \ldots, d_n)$ where $d_i \in \{-1, 1\}$ such that

$$d_1 r_1(u_1 x + v_1 y) + \cdots + d_n r_n(u_n x + v_n y) = 0$$
$$\wedge$$
$$d_1 r_1(u'_1 x + v'_1 y) + \cdots + d_n r_n(u'_n x + v'_n y) \neq 0.$$

Define $c = (c_1, \ldots, c_n)$ such that $c_i = d_i r_i$. Clearly, $c \in A(f)$ and $c \notin A(f')$, thus $\exists c \in A(f) \ominus A(f')$. [3] Recall that $r_i \leq k_i$ and $\sum_i^n k_i = k$ and $k \leq \hat{k}$, hence $\sum_i^n |c_i| \leq \hat{k}$ and we have proven the contrapositive statement. ∎

Corollary 3.2 *If $A(f) \ominus A(f') = \emptyset$ then for all k the moments $M_k(f) = M_k(f')$.*

Proof This must follow directly from the theorem to avoid a contradiction. ∎

Theorem 3.1 relates the histogram to the integer frequencies through the solution set of a pair of Diophantine equations. Given any pair of functions in $F(w)$ the number of identical moments are found by generating the solutions to the Diophantine equations until a member of $A(f) \ominus A(f')$ is found. Note that this process will always terminate provided $A(f) \neq A(f')$, which is the case if (u, v) and (u', v') are linearly independent. From the corollary, if the frequencies are related by a linear transform then all of the moments of f and f' will be the same and $p(f) = p(f')$. By relaxing the constraint that all of the moments must match, functions can be generated that are not related by an affine transformation yet have $p(f) \approx p(f')$, which in the presence of quantization and noise will be statistically indistinguishable.

4. Moments of Filtered Images

Next, we consider the marginal distributions of the subband images of f and f'.

Theorem 4.1 *Let f and f' be functions in $F(w)$. Define $\hat{k}$ to be $< \sum_i^n |c_i|$ for all $c \in A(f) \ominus A(f')$. Let g be any symmetric or anti-symmetric function with the Fourier transform G. If $\forall i, G(u_i, v_i) = G(u'_i, v'_i)$ then $\forall k : k \leq \hat{k}, M_k(g * f) = M_k(g * f')$.*

Proof Case 1: The function g is symmetric. The symmetry implies that G is real valued, hence from the properties of the Fourier transform it can be shown that

$$\begin{aligned} g * f &= \sum_{i=1}^{n} G(u_i, v_i) w_i \cos(ux + vy) \\ g * f' &= \sum_{i=1}^{n} G(u'_i, v'_i) w_i \cos(u'x + v'y). \end{aligned}$$

If $G(u_i, v_i) = G(u'_i, v'_i)$ define $m = (m_1, \ldots, m_n) : m_i = G(u_i, v_i) w_i$, thus $g * f \in F(m)$ and $g * f' \in F(m)$. Note that the integer frequencies have remained the same therefore $A(g * f) = A(f)$ and $A(g * f') = A(f')$. This implies that $\hat{k} < \sum_i^n |c_i|$ for all $c \in A(g * f) \ominus A(g * f')$ thus $\forall k : k \leq \hat{k}, M_k(g * f) = M_k(g * f')$ follows from theorem 3.1.

Case 2: The function g is anti-symmetric. The anti-symmetry implies that G is imaginary, hence from the properties of the Fourier transform it can be shown that

$$\begin{aligned} g * f &= \sum_{i=1}^{n} G(u_i, v_i) w_i \sin(ux + vy) \\ g * f' &= \sum_{i=1}^{n} G(u'_i, v'_i) w_i \sin(u'x + v'y). \end{aligned}$$

The sine functions are cosine functions with a $\pi/2$ shift. Moments are invariant to constant shifts hence case 2 follows from case 1. ∎

Together, theorems 3.1 and 4.1 provide sufficient conditions for ensuring the first k moments of the histograms of a set of subband images match for a pair of functions f and f' and a set of filters $g = \{g_1, \ldots\}$. We are not concerned with the actual value of k except to note that it can always be defined large enough such that the moments above k are lost in quantization and noise thus we claim $p_\alpha(f) = p_\alpha(f'), \forall \alpha : g_\alpha \in g$. To construct examples of f and f' with matching moments, the frequencies (u, v) and (u', v') must be chosen carefully such that:

1. The solutions to the Diophantine equations (8) of magnitude $\leq k$ are identical.
2. For each filter $g_\alpha \in g$ with Fourier transform G_α we have $\forall i, G_\alpha(u_i, v_i) = G_\alpha(u'_i, v'_i)$.

The general approach taken to construct examples of f and f' is to define two sets of frequencies $u_a = \{(u_1, v_1), \ldots, (u_n, v_n)\}$ and $u_b = \{(u_1, v_1), \ldots, (u_m, v_m)\}$. Then define $f = f_a + f_b$ where f_a is constructed from the frequencies u_a and f_b from u_b. The weights $(w_1, \ldots, w_n, w_{n+1}, \ldots, w_m)$ are set to 1. Similarly, $f' = f_a + R(f_b)$ is defined where R is a rotation of the frequencies u_b and the weights are also set to 1. The exact choice of R will depend on the filters which is addressed later. The key is to choose the frequencies such that u_a, u_b do not "interact" and $u_a, R(u_b)$ do not "interact" to contribute to any of the moments $\leq k$. Formally, if $c = (c_1, \ldots, c_{n+m}) \in A(f) : \sum |c_i| \leq k$ then we require $(c_1, \ldots, c_n) \in A(f_a) \vee (c_{n+1}, \ldots, c_m) \in A(f_b)$. Likewise, if $c \in A(f') : \sum |c_i| \leq k$ then we require $(c_1, \ldots, c_n) \in A(f_a) \vee (c_{n+1}, \ldots, c_m) \in A(R(f_b))$. These properties imply that $c \notin A(f) \ominus A(f')$. Rather than iterating through possible choices of u_a and u_b simple

[3] $\ominus$ is the symmetric set difference operator, which is the elements belonging to one but not both sets.

Proceedings of the Ninth IEEE International Conference on Computer Vision (ICCV 2003) 2-Volume Set
0-7695-1950-4/03 $17.00 © 2003 IEEE

geometric heuristics are applied to avoid interacting with any small moments. For example, ensuring that no two members of u_b and $R(u_b)$ are parallel to any two members of u_a . The frequencies are also chosen such that "visually distinct" patterns emerge. This is accomplished by placing members of u_b close to members of u_a thus creating interference patterns that will not exist between $R(u_b)$ and u_a. Note that although R is an affine transform the functions f and f' will in general not be related by an affine transformation. To handle certain classes of filters we need one more theorem.

Theorem 4.2 *For a function $f \in F(w)$, if $\forall i,\ u_i$ are odd or $\forall i,\ v_i$ are odd then $\forall k$ such that k is odd,$M_k(f) = 0$.*

Proof Consider an arbitrary odd moment k and an arbitrary κ and r in the summation of equation (6). To prove that $M_k = 0$ we only need to show that $t(u, v, r)$ is zero for the arbitrary term.

Define the set O_κ to contain the odd parity terms of κ. Define the set O_r to contain the odd parity terms of r. Define the set O_u to contain the odd parity terms of $d_i r_i u_i$.

The fact that k is odd and $\sum_i^n k_i = k$ implies that an odd number of odd parity k_i terms must exist, hence $|O_\kappa|$ is odd. We need only consider the terms in (6) where $\mathcal{C}(r, \kappa) \neq 0$. Hence, from (5) we also know that $|O_r|$ is odd. Now, if all of the u_i are odd, the fact that $|O_r|$ is odd implies that $|O_u|$ is odd. If $|O_u|$ is odd then $\sum_i^n d_i r_i u_i \neq 0$, hence $t(u, v, r) = 0$ for the arbitrary term and therefore $M_k = 0$. The same holds true if the v_i are odd. ∎

A filtering operation does not alter the frequencies of a function it only effects the weights. Thus, by using odd parity frequencies we have a way of ensuring all of the odd moments are zero for all of the subband images of f and f'.

Now we are ready to consider examples of f and f' for several classes of commonly used filters. In order to create example patterns discrete images must be formed. However, the theorems have assumed a continuous image model and will not necessarily be true for sampled functions. We work around this problem by randomly sampling the functions inside each pixel in which case the expected value of the first k moments match.

4.1. Rotationally Symmetric Filters

Rotationally symmetric filters include Gaussian, Laplacian of Gaussian and other center surround kernels. These filters are defined with a scale parameter σ, which determines the size of the kernel. The common feature of this class of filters is a rotationally symmetric Fourier transform. Any rotation R that maps the integer frequencies to another set of integer frequencies results in exactly the same Fourier coefficients. Integer rotations are easily be achieved by interchanging the u and v frequencies. Thus, $f(u, v) = f_a(u, v) + f_b(u, v)$ and $f'(u, v) = f_a(u, v) + f_b(v, u)$ have the same weights after convolution with any of these filters applied at any scale even if an infinite number of scales is used. Figures (a) and (b) show examples of such patterns sampled with 256×256 pixels. Next to each pattern is a figure showing the distribution of the frequencies used to construct the pattern. In (a) the fixed frequencies u_a are denoted by circles and the rotated frequencies u_b are denoted by plus signs and the first 35 moments are identical. Several rotations of the frequencies are applied to (b) and the first 4 moments are identical. Each pattern is convolved with both Gaussian and Laplacian of Gaussian kernels at scales $\sigma = \{1, 2, \ldots, 8\}$ for a total of 16 subband images. For each pair of subband images, histograms are formed with 100 bins and the L_1 norm is computed. For all histograms the difference is between 1 and 3 percent of the number of pixels, a negligible amount. Next to the patterns in (a) and (b) are example histograms for the Gaussian and Laplacian filters that resulted in largest L_1 norm.

4.2. Derivative Filters

Another commonly used class of filters are the derivatives of a Gaussian function. They have 3 parameters, the order of the spatial derivative in the x and y direction and the scale of the Gaussian. Even order derivative filters are symmetric while odd order derivative filters are anti-symmetric. The Fourier transform of a derivative filter is a polynomial in u, v modulated by a Gaussian. It is straightforward to show that the coefficients of a derivative of Gaussian are such that $G(u, v) = G(-u, v)$ for symmetric filters and $G(u, v) = -G(-u, v)$ for anti-symmetric filters. A reflection transformation preserves the coefficients of a symmetric filter, thus $f(u, v) = f_a(u, v) + f_b(u, v)$ and $f'(u, v) = f_a(u, v) + f_b(-u, v)$. To handle anti-symmetric filters an additional constraint is needed. For an anti-symmetric derivative filter g we have $g * f = g * f_a(u, v) + g * f_b(u, v)$ and $g * f' = g * f_a(u, v) - g * f_b(-u, v)$. This causes a change of sign in the odd moments, resulting in a reflected histogram. However, if all of the odd moments are zero then the histogram is symmetric and invariant to a reflection. By theorem 4.2 imposing the constraint that either the u or v frequencies of f_b are odd forces the odd moments to zero. Example patterns are shown in (c) and (d). For (c) the first 9 moments match and for (d) the first 39 moments match. A total of 80 subband images are formed using derivative filters of all orders between 0 and 3 and each filter is applied at scales $\sigma = \{1, 2, \ldots, 8\}$. Again all histograms matched within an L_1 difference of 1 to 3 percent. The histogram with the largest L_1 norm and the histogram of the unfiltered image are shown.

4.3. Gabor Filters

Gabor functions are a class of oriented filters and are extensively used in texture modeling. These functions come in symmetric/anti-symmetric pairs where each pair is a cosine/sine wave modulated by a Gaussian function. The

Proceedings of the Ninth IEEE International Conference on Computer Vision (ICCV 2003) 2-Volume Set
0-7695-1950-4/03 $17.00 © 2003 IEEE

Fourier transform of a Gabor function is a Gaussian function centered at the frequency of the cosine/sine wave. A bank of Gabor filters is usually designed to tile the frequency plane with Gaussian functions occurring at regular intervals in orientation and scale. Given a bank of Gabor filters with orientations $\{\theta_1, \theta_2, \ldots\}$ and a spacing of ψ between consecutive orientations, the Fourier coefficients of all of the filters at orientation θ_i will have the following property

$$G(r\cos\theta_i+\frac{\psi}{2}, r\sin\theta_i+\frac{\psi}{2}) = G(r\cos\theta_i-\frac{\psi}{2}, r\sin\theta_i-\frac{\psi}{2}).$$

Therefore, if we define the frequencies of f_b such that they have the orientations $\{\psi/2, 5\psi/2, 9\psi/2, \ldots\}$ and R is a rotation of angle ψ the coefficients of f_b will be identical to $R(f_b)$. An example of such a pattern is discussed in the following subsection.

4.4. Combining Filters

Function pairs that share the same marginal statistics for combinations of these filters also exist. The patterns derived for derivative and Gabor filters also have the same distributions for rotationally symmetric filters. However, the transformation R for derivative filters is a reflection which will not satisfy the constraint on R for Gabor filters. But the rotation for Gabor filters is also a reflection, thus the constraints for Gabor filters imply the constraints for the symmetric derivative filters. To satisfy the constraints for the anti-symmetric derivative filters the odd frequency constraint must be added to the Gabor filters. Figure (e) shows an example of a pair of patterns that have been defined this way for a bank of Gabor filters at orientations $\{0, \pi/6, \pi/3, \ldots\}$. The patterns only share the first 5 moments however they still have nearly identical histograms. Each pattern has been convolved with a total of 48 Gabor filters (6 orientations and 4 scales), 80 derivative filters and 16 rotationally symmetric filters. The marginal densities of all 144 subband images matched to within 6 percent. A sampling of the histograms is shown in (e).

5. Summary and Conclusions

To summarize, we have shown the existence of a class of patterns that even if Gaussian, Laplacian, derivative and Gabor filters are applied at an *infinite* number different scales the patterns will not be distinguishable from the marginal statistics alone. The main implication of this result is that some form of joint statistical information should be added if we wish to have vision systems that can discriminate among a class of patterns at least as large as those that are distinct to a human observer. It seems that in order to avoid joint statistics oriented filters tuned to very fine degrees would have to be employed. Thus we are faced with the question of efficiency. Clearly, the marginal statistics contain a large amount of information about a visual pattern. We would like to determine the minimal amount of joint statistical information that should be added.

References

[1] M. Mandal and T. Aboulnasr, "Image indexing using moments and wavelets," *IEEE Transactions on Consumer Electronics*, vol. 42, no. 3, pp. 557–565, 1996.

[2] B. Schiele and J. Crowley, "Recognition without correspondence using multidimensional receptive field histograms," *International Journal of Computer Vision*, vol. 36, no. 1, 2000.

[3] C. Schmid, "Constructing models for content-based image retrieval," *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*, 2001.

[4] H. Schneiderman and T. Kanade, "A statistical method for 3d object detection applied to faces and cars," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2000.

[5] B. Julesz, "Visual pattern discrimination," *Transactions of Information Theory*, , no. 8, pp. 84–92, 1962.

[6] T. Caelli and B. Julesz, "On perceptual analyzers underlying visual texture discrimination," *Part I: Biological Cybernetics*, vol. 28, pp. 167–175, 1978.

[7] G. Cross and A. Jain, "Markov random field texture models," *The IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 5, pp. 25–39, 1983.

[8] R. De Valois M. Silverman, D. Grosof and S. Elfar, "Spatial-frequency organization in primate striate cortex," *Proc. of the Natl. Acad. Sci.*, 1989.

[9] C. Chubb and M.S. Landy, "Orthogonal distribution analysis: A new approach to the study of texture perception," in *Computational Models of Visual Processing*, M.S. Landy, Ed. MIT Press, 1991.

[10] J.R. Bergen and E.H. Adelson, "Theories of visual texture," in *Spatial Vision*, D. Regan, Ed. CRC Press, 1991.

[11] D. Heeger and J Bergen, "Pyramid-based texture analysis/synthesis," in *Proceedings of SIGGRAPH*, 1995.

[12] Y. Wu S.C. Zhu and D. Mumford, "Filters, random-fields and maximum-entropy (frame): Towards a unified theory for texture modeling," *International Journal of Computer Vision*, vol. 27, no. 2, pp. 107–126, 1998.

[13] X. Liu S.C. Zhu and Y. Wu, "Exploring texture ensembles by efficient markov chain monte carlo," *The IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 6, 2000.

[14] J. De Bonet and P. Viola, "A non-parametric multi-scale statistical model for natural images," *Advances in Neural Information Processing*, vol. 10, 1997.

[15] J. Portilla and E. Simoncelli, "A parametric texture model based on joint statistics of complex wavelet coefficients," *International Journal of Computer Vision*, vol. 40, no. 1, pp. 49–71, 2000.

[16] T. Leung and J. Malik, "Representing and recognizing the visual appearance of materials using three-dimensional textons," *International Journal of Comptuer Vision*, Dec 1999.

[17] K. Dana and S. Nayar, "Histogram model for 3d textures," in *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*, 1998.

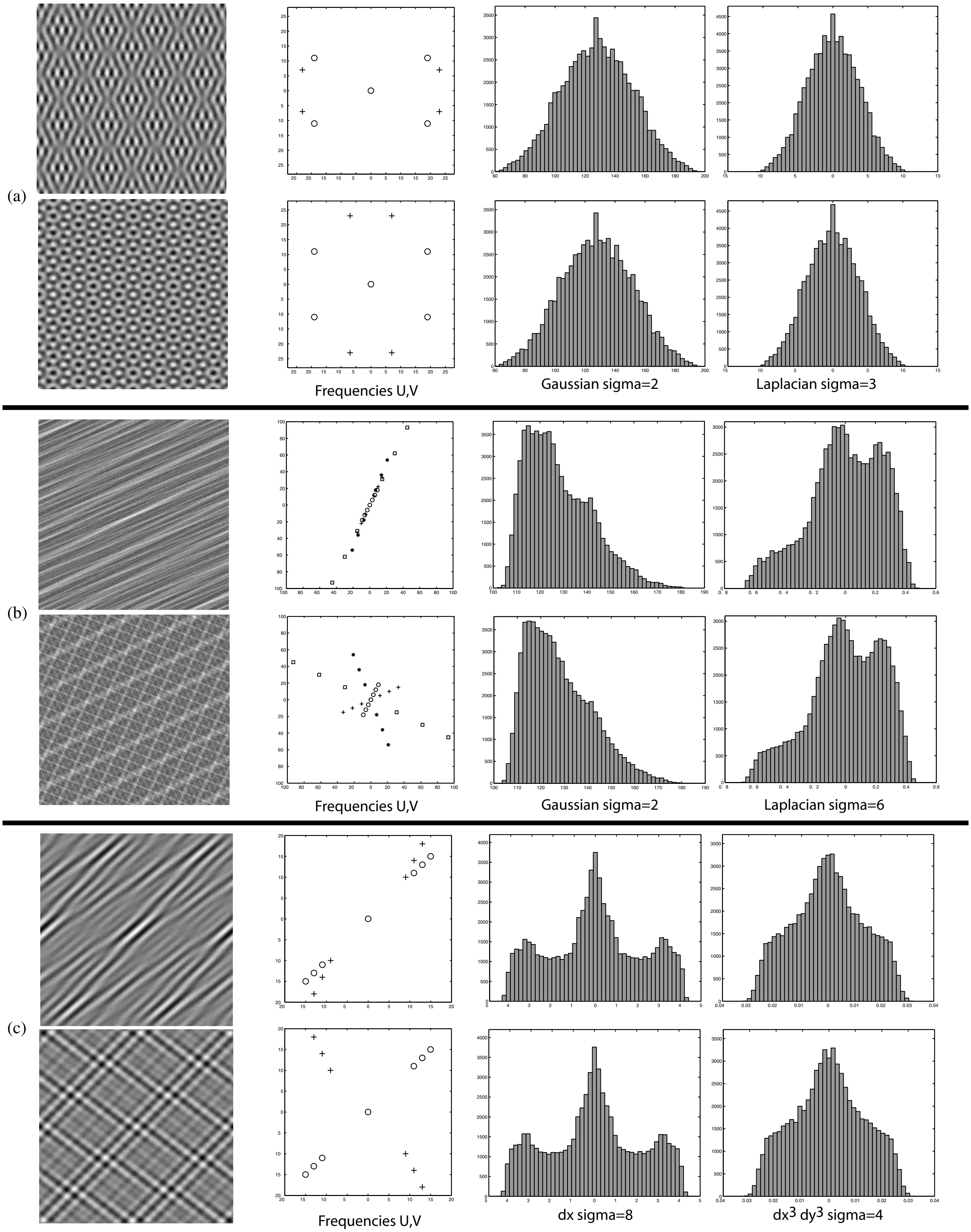
[18] M. Varma and A. Zisserman, "Classifying images of materials: achieving viewpoint and illumination independence," in *Proceedings of European Conference on Computer Visions*, 2002.

[19] M. Grossberg E. Hadjidemetriou and S. Nayar, "Histogram preserving image transformations," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2000.

[20] M. Grossberg E. Hadjidemetriou and S. Nayar, "Spatial information in multiresolution histograms," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2001.

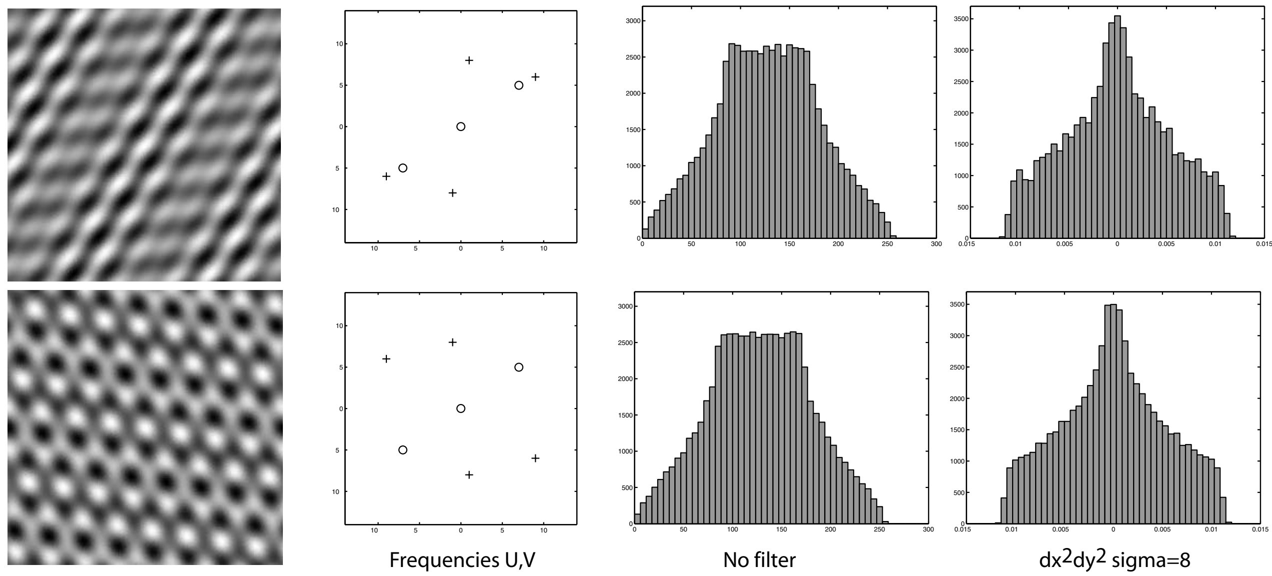
[21] P. Meyer, *Introductory Probability and Statistical Applications*, Addison Wesley, 2nd edition, 1970.

Proceedings of the Ninth IEEE International Conference on Computer Vision (ICCV 2003) 2-Volume Set
0-7695-1950-4/03 $17.00 © 2003 IEEE

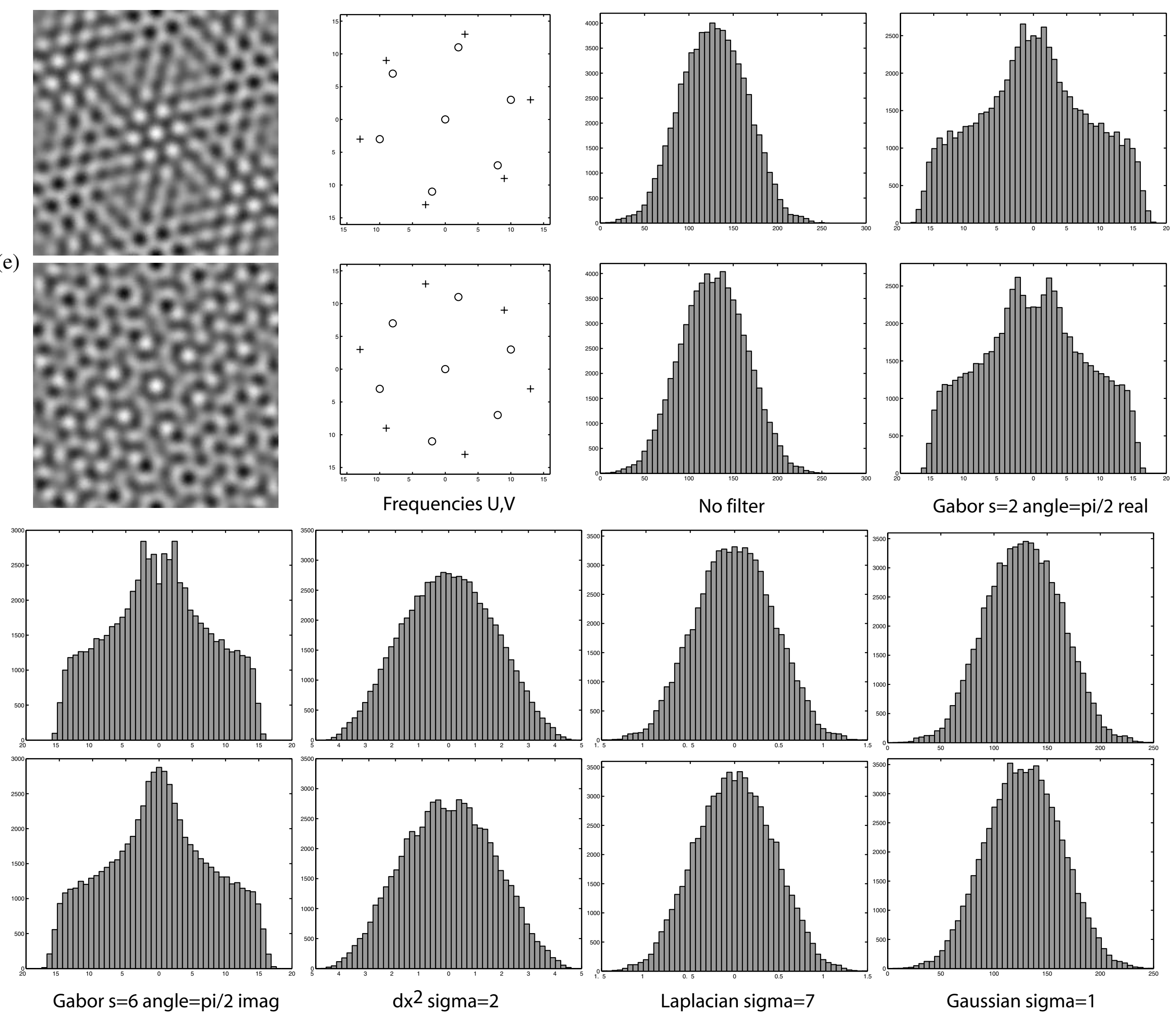


Proceedings of the Ninth IEEE International Conference on Computer Vision (ICCV 2003) 2-Volume Set
0-7695-1950-4/03 $17.00 © 2003 IEEE

(d)



(e)



Proceedings of the Ninth IEEE International Conference on Computer Vision (ICCV 2003) 2-Volume Set
0-7695-1950-4/03 $17.00 © 2003 IEEE