Fig. 7. Conditional information of sentences $x \in L(G_F)$ given independent sentences $4 = abab$ and $27 = bbabaa$.

tional information relative to these pairs to obtain an internal grouping of the sentences of the given grammar.

VI. Summary and Conclusions

In this correspondence we have considered the complexity of sentences with emphasis on the role that conditioning plays in definitions of complexity. Based on this definition we have shown that it is possible to construct marginal, conditional, and mutual syntactical information measures which satisfy some basic properties that makes them useful in applications. It has been demonstrated that the notions of complexity and information can be combined to yield a consistent description of the structure of sentences. Also, since the concept of universality was not necessary the results obtained are easier to apply than the general theory of complexity permits.

References

[1] A. V. Aho and T. G. Peterson, "A minimum distance error-correcting parser for context-free languages," *SIAM J. Comput.*, vol. 1, pp. 305–312, 1972.

[2] G. J. Chaitin, "A theory of program size formally identical to information theory," *J. Ass. Comput. Mach.*, vol. 22, no. 3, pp. 329–340, 1975.

[3] K. S. Fu, *Syntactic Methods in Pattern Recognition.* New York: Academic, 1974.

[4] L. W. Fung and K. S. Fu, "Stochastic syntactic decoding for pattern classification," *IEEE Trans. Comput.*, vol. C-24, no. 6, pp. 662–667, 1975.

[5] A. N. Kolmogorov, "Three approaches to the definition of the concept of the amount of information," *Probl. Inform. Transm.*, vol. 1, pp. 3–11, 1965.

[6] R. A. Kraak, D. E. Boekee, and E. Backer, "Complexity and information content of sentences generated by certain pushdown machines," Dept. Elec. Eng., Delft Univ. Technol., The Netherlands, Rep. IT-80-06, 1980.

[7] S. K. Leung-Yan-Cheong and T. M. Cover, "Some equivalences between Shannon entropy and Kolomogorov complexity," *IEEE Trans. Inform. Theory*, vol. IT-24, no. 3, pp. 331–339, 1978.

[8] S. Y. Lu, "A tree-to-tree distance and its applications to cluster analysis," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-1, no. 2, pp. 219–224, 1979.

[9] S. Y. Lu and K. S. Fu, "A sentence to sentence clustering procedure for pattern analysis," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-8, no. 5, pp. 381–389, 1978.

[10] R. J. Solomonoff, "A formal theory of inductive inference," *Inform. Contr.*, vol. 7, pp. 1–22 and 224–254, 1964.

[11] R. A. Wagner and M. J. Fisher, "The string-to-string correction problem," *J. Ass. Comput. Mach.*, vol. 21, no. 1, pp. 168–173, 1974.

Multispectral Texture

AZRIEL ROSENFELD, FELLOW, IEEE, CHENG-YE WANG, AND ANGELA Y. WU

Abstract—Textures in single-band images are often characterized by statistics of the joint distributions of pairs of gray levels for pairs of pixels in given relative positions, or by statistics of absolute gray level differences for such pairs of pixels. Joint distributions of pairs of spectral vectors in multiband images are cumbersome—since for k bands they are $2k$-dimensional—but absolute difference distributions are less so, e.g., for two bands they are only two-dimensional. The possibility is discussed of using statistics of absolute difference distributions for characterizing textures in multiband images, with emphasis on the two-band case.

I. Introduction

Many different types of features have been used for texture analysis and classification; see [1] for a recent review. Essentially all of this work has dealt with single-band images rather than with color or multispectral images. When texture analysis is used for multispectral imagery, it is applied to a single band (possibly a composite of the original bands, or an "eigenband" resulting from a Karhunen–Loève transformation), and if desired, the results are treated as an additional "texture band"; but texture features are not commonly, if ever, measured for multiband data directly. The purpose of this correspondence is to introduce a class of texture features that are defined for multiband imagery and that are computationally quite tractable in the two-band case.

The particular class of texture features which we will generalize to the multiband case are statistical features derived from pairs of pixels in given relative positions. For single-band images, the joint distribution of the gray levels of such a pair of pixels can be represented by a "co-occurrence matrix" which tabulates how often each possible pair of gray levels occurs in the image in the given relative position, and we can define texture features by computing various statistics from such matrices (e.g., moment of inertia about the main diagonal, entropy, etc). This concept generalizes immediately to multi-band images but is computationally cumbersome; even in the two-band case, the joint distribution of pairs of two-vectors in a given relative position requires a four-dimensional matrix for its representation, which is expensive in storage space unless the values in the bands are very coarsely quantized. (It might be possible, in principle, to use sparse matrix techniques to handle high-dimensional co-occurrence matrices; but we shall not pursue this possibility here.)

An alternative to using joint distributions of pairs of pixel gray levels in given relative positions is to use only the distribution of absolute differences of such pairs of gray levels. It was seen in [2] that for some texture classification tasks, features based on such distributions are just as effective as features derived from joint gray level distributions. In the single-band case, an absolute difference distribution is represented by a (one-dimensional) histogram of the absolute differences, and we can define texture features by computing various statistics from such histograms, e.g., their means, variances, entropies, etc. In the two-band case, it would be represented by a two-dimensional scatter plot show-

ing how often each (difference in band 1, difference in band 2) pair occurs for pixel pairs in the given relative position. Thus for small numbers of bands (two, especially), texture analysis based on absolute difference statistics is computationally quite tractable.

Section II of this correspondence defines a class of multiband texture features based on absolute difference statistics, and Section III gives examples of results obtained when these features are computed for some simple two-band textures.

II. FEATURES

A. Single-Band Features

Let us first briefly review the definitions of co-occurrence matrices and absolute difference histograms for single-band images. Let $\delta \equiv (\Delta x, \Delta y)$ be a relative position vector, and let the gray levels of the given images be $0, 1, \cdots, m - 1$. The *co-occurrence matrix* M_δ is an $m \times m$ matrix whose (i, j)th element is the number of pairs of pixels in relative position δ that have the pair of gray levels (i, j). For a uniformly textured image having a given gray level probability density, concentration of high values near the main diagonal of M_δ suggests that the texture is composed of uniform patches that are large relative to $|\delta|$ (implying that two gray levels δ apart tend to be similar). Thus the moment of inertia of M_δ about its main diagonal is a measure of the "busyness" of the texture relative to $|\delta|$. This is one simple example of how statistics computed from M_δ, for various δ's, can provide information about the nature of the texture.

Similarly, the *difference histogram* D_δ is an m-vector whose kth element is the number of pairs of pixels in relative position δ that have absolute gray level difference k. Note that if we sum M_δ along lines parallel to its main diagonal, we obtain a difference histogram, but for signed rather than absolute differences; to obtain D_δ, we need only make M_δ symmetric by adding pairs of elements symmetric with respect to the main diagonal, and then summing the upper triangle of M_δ along lines parallel to the diagonal. Thus the moment of inertia of D_δ around the origin ($k = 0$) is the same as the moment of inertia of M_δ around the main diagonal, indicating that D_δ too can be used as a source of texture features.

Co-occurrence matrices and difference histograms are commonly computed for images whose gray level probability densities have been standardized, e.g., by histogram flattening. If this were not done, contrast effects would be confused with coarseness effects; if we increase the contrast of an image, its M_δ entries are spread outward from the main diagonal and its D_δ entries from the origin.

B. Multiband Features

The values of the pixels in a b-band image are b-vectors of the form $(Z_1, \cdots, Z_b)$, $0 \le Z_h \le m - 1$, $1 \le h \le b$. Thus the b-band analog of a co-occurrence matrix M_δ is a $2b$-dimensional $m \times m$ $\cdots \times m$ array whose $(i_1, \cdots, i_{2b})$th element is the number of pairs of pixels in relative position δ that have the pair of b-vectors $((i_1, \cdots, i_b), (i_{b+1}, \cdots, i_{2b}))$ as values. Evidently, even for small values of m and b, such an M_δ is cumbersome to work with, e.g., for $m = 8$ and $b = 2$, it has $m^{2b} = 8^4 = 2^{12} = 4096$ elements, and this number grows rapidly with both m and b.

The situation is somewhat more manageable if we work with difference histograms rather than co-occurrence matrices. A b-band difference histogram D_δ is a b-dimensional $m \times m \cdots \times m$ array whose $(k_1, \cdots, k_b)$th element is the number of pairs of pixels in relative position δ that have the b-vector of absolute differences $(k_1, \cdots, k_b)$ in bands $1, \cdots, b$, respectively. The size of D_δ for small values of b is quite manageable; e.g., for $m = 8$ and $b = 2$ it has only $m^b = 64$ elements.

Some insight into the possible forms of multiband scatter plots (again, for simplicity we assume $b = 2$) can be obtained by considering two simple hypothetical examples.

1) Suppose that the texture is composed of small patches on a background, where the patches and the background have a greater reflectivity difference in one band than in the other. For a given δ, smaller than the average patch size or spacing, the difference histogram in each band is a mixture of within-patch and within-background differences (presumably near 0) and patch–background differences (larger). In this case, the two-dimensional D_δ scatter plot should consist of a cluster near $(0, 0)$ and a cluster near (d_1, d_2), where d_1 and d_2 are the expected patch–background differences in the two bands.

2) Suppose that the texture arises from an undulating surface in which slope differences give rise to intensity differences in the image, and the change in intensity as a function of slope is different for the two bands. A given displacement δ corresponds to a given expected slope difference, hence to a pair of expected intensity differences in the two bands.

Note that in both of these cases, the differences in the two bands are quite correlated.

What types of statistics would it be useful to measure for a b-band M_δ or D_δ? (We shall assume, for convenience, that the probability densities of values have been standardized, e.g., by histogram flattening of each band.) Evidently, the spread of values relative to the main diagonal or origin is still relevant, but we should be able to analyze it in greater detail, since it has more degrees of freedom. For simplicity, let us consider only D_δ and only the two-band case. In this case D_δ is a two-dimensional array whose (i, j) element is the number of pairs of pixels in relative position δ that have absolute difference i in the first band and j in the second band. We can make the following qualitative observations about such an array.

a) The spread of values away from the origin is a measure of texture "busyness," since high values far from the origin imply the frequent occurrence of high absolute differences in one or both bands.

b) The spread of values away from the main diagonal (as measured, e.g., by the moment of inertia of D_δ about the diagonal) is a measure of relative texture "busyness" in the two bands; high values far from the diagonal imply many cases where one absolute difference is quite different from the other.

c) The asymmetry of the values relative to the main diagonal (as measured, e.g., by the slope of the principal axis of D_δ relative to the diagonal direction) indicates which of the two bands is "busier."

Analogous remarks can be made about the b-band case for $b > 2$.

Evidently, measures such as b) and c), can only be obtained from a two-band scatter plot such as D_δ; they could not be derived by analyzing the two bands separately, since they measure the correlatedness of the absolute difference values between the bands. Thus it is clear that texture features based on a two-band D_δ can provide information about the texture not available from single-band texture features. This means that, in principle, pairs of two-band textures exist that can be discriminated easily when features based on a two-band D_δ are used but that are hard to discriminate based on single-band features. Of course, this does not imply that such pairs of textures will be very common; it may be difficult to find such examples.
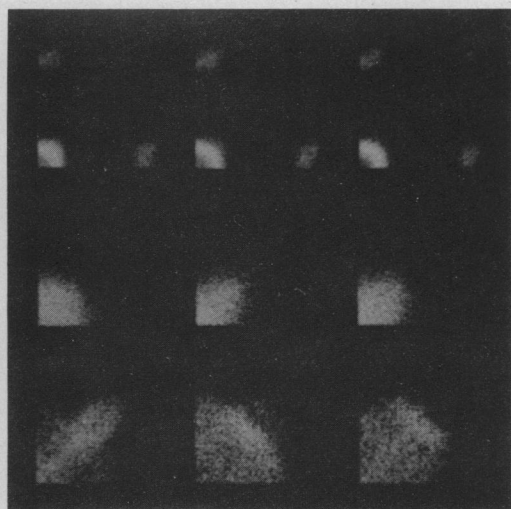
III. EXAMPLES

The following examples are based on a class of synthetic single-band textures derived from stationary random field models having given types of neighbor dependence [3]. Three examples of textures generated by such models are shown in Fig. 1.

Artificial two-band textures were created from these single-band textures as follows: let T be a given single-band texture, and let T_i be the two-band texture whose bands are T and T shifted by the

Fig. 1.   Three textures generated by stationary random field models [3].

TABLE I
ASM AND CON Statistics for the D_δ Plots in Fig. 2

Relative Texture	Displacement of "bands"	ASM $\times 10^3$	CON
(a)	1	4.0	530
	2	4.4	532
	4	4.6	458
(b)	1	2.1	103
	2	2.4	64
	4	2.2	75
(c)	1	1.1	123
	2	1.0	295
	4	1.0	198

amount δ_i. The smaller $|\delta_i|$ (relative to the neighborhood used in defining T), the more these two bands should be correlated. Thus if we consider a set of two-band textures T_i having $|\delta_i|$ of various sizes, we should obtain rather different two-band scatter plots D_δ when $|\delta|$ is small.

Fig. 2 shows examples of the D_δ plots obtained for the synthetic textures in Fig. 1, using $\delta_1 = (1,0)$, $\delta_2 = (2,0)$, $\delta_3 = (4,0)$, and $\delta = (1,0)$. We see that the scatter plots are indeed rather different for the different δ_i. To quantify this difference, Table I shows the values of two statistics measured for these scatter plots: the sum of the squares of the values (ASM) and the moment of inertia about the 45° diagonal (CON) [1], [2].

Multiband features should be especially useful if we apply them to "eigenbands" produced by applying a principal-components transformation to the original bands, since the eigenbands should be less correlated than the originals. On eigenbands and their role in the analysis of multispectral imagery see [4].

We now give an example in which two-band texture features, derived from eigenbands, appear to provide better information than is available from the individual bands. Figs. 3–6 show a set of 32 64 × 64 pixel images of the Copper Mountain, WY, area. The images in Figs. 3 and 5 are in eigenband 1, while those in Figs. 4 and 6 are eigenband 3. The original images were classified by a geologist into two categories: uranium-bearing (U) and nonuranium bearing (NU); these classes are shown in Figs. 3, 4 and 5, 6, respectively.

In earlier experiments [5], texture features derived from the eigenbands gave partial separations between the U and NU samples. Tables II–IV show, for eigenband 1, eigenband 3, and the pair of bands, the values of the texture features ASM, CON, and IDM. [For a scatter plot (e_{ij}), we have ASM $\equiv \Sigma\Sigma e_{ij}^2$,

Fig. 2. D_δ plots for two-band textures derived from the textures in Fig. 1. One band is cyclically shifted relative to the other by 1, 2, or 4 pixels to produce the plots in left, center, and right columns, respectively.
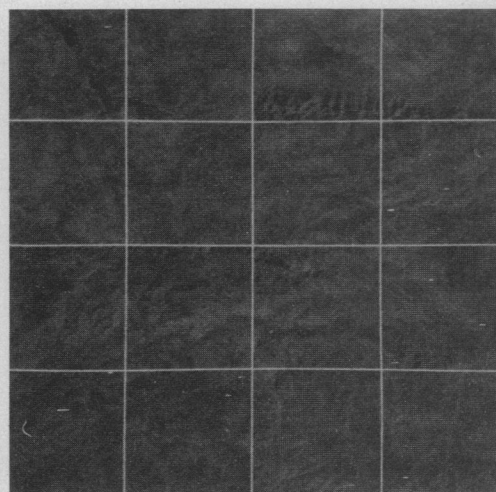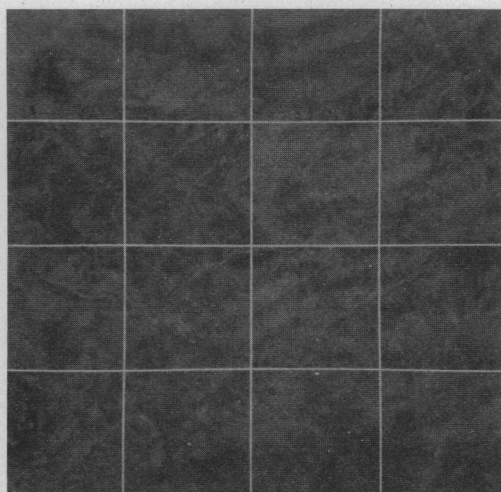

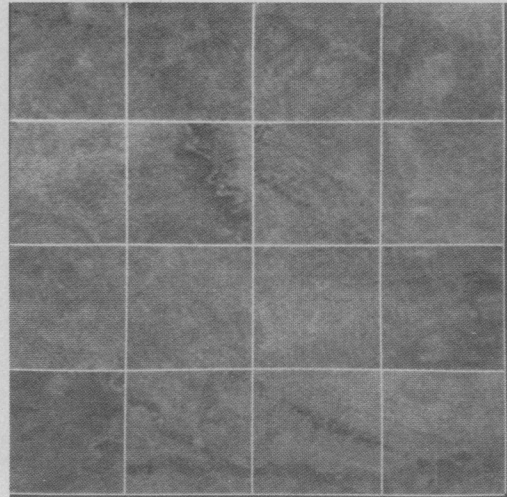Fig. 3.   Uranium-bearing images, eigenband 1.
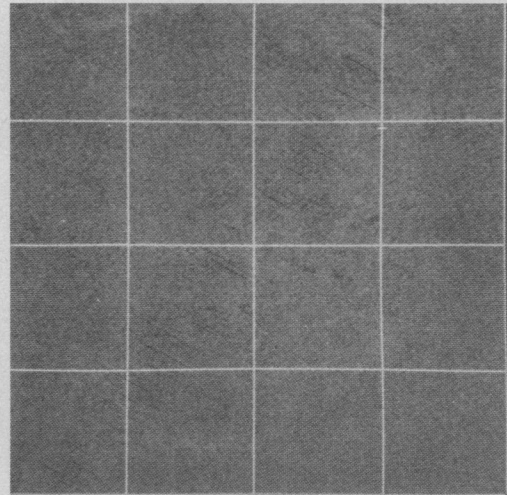
Fig. 4.   Nonuranium-bearing images, eigenband 1.
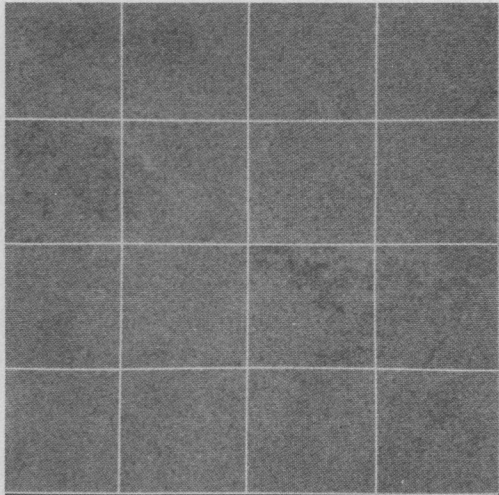

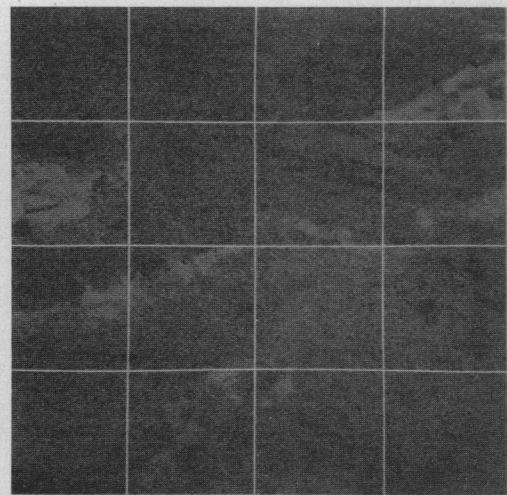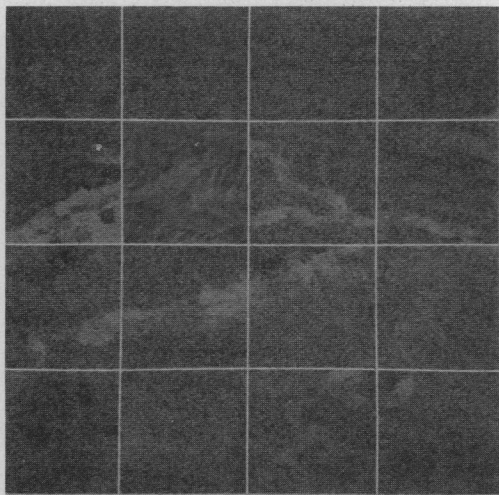Fig. 5.   Uranium-bearing images, eigenband 3.

Fig. 6.   Nonuranium-bearing images, eigenband 3.

TABLE II
VALUES OF TEXTURE FEATURES DERIVED FROM EIGENBAND 1

Uranium image no.	CON	IDM	ASM	Non-uranium image no.	CON	IDM	ASM
1	43	248	397	1	27	326	1146
2	39	256	445	2	29	332	1037
3	36	241	553	3	32	302	758
4	47	213	436	4	42	254	558
5	38	248	587	5	52	216	455
6	38	269	668	6	45	240	516
7	44	241	419	7	44	244	319
8	37	244	477	8	33	297	782
9	36	241	541	9	33	296	879
10	37	246	584	10	26	310	950
11	43	221	440	11	30	304	750
12	37	237	483	12	40	267	491
13	52	207	338	13	52	226	390
14	37	252	584	14	42	275	486
15	32	271	639	15	36	261	720
16	35	260	599	16	36	254	545
17	42	229	423	17	36	266	657
18	39	232	515	18	34	269	670
19	70	189	257	19	42	229	441
20	46	236	492	20	35	285	660
21	30	284	889	21	46	230	399
22	33	262	691	22	53	212	313
23	41	235	550	23	48	233	340
24	45	226	478	24	43	250	503
25	44	217	455	25	43	260	529
26	41	243	647	26	46	233	369
27	32	267	792	27	56	198	307
28	28	280	810	28	45	237	489
29	28	286	818	29	55	232	344
30	46	208	470	30	50	213	463
31	36	244	662	31	56	208	424
32	49	212	471	32	44	219	443

TABLE III
VALUES OF TEXTURE FEATURES DERIVED FROM EIGENBAND 3

Uranium image no.	CON	IDM	ASM	Non-uranium image no.	CON	IDM	ASM
1	121	118	201	1	118	126	249
2	117	121	244	2	121	122	224
3	116	120	261	3	116	123	253
4	122	116	250	4	124	120	227
5	120	121	246	5	130	116	232
6	111	123	272	6	122	119	231
7	123	119	226	7	129	112	184
8	125	117	235	8	124	121	209
9	124	114	236	9	111	126	260
10	118	119	265	10	114	124	250
11	125	115	198	11	115	127	262
12	125	112	221	12	115	123	259
13	128	121	208	13	125	115	224
14	118	119	242	14	121	117	223
15	119	117	255	15	121	123	239
16	119	121	258	16	121	121	218
17	124	120	240	17	119	123	237
18	128	116	230	18	121	118	245
19	132	115	210	19	125	115	212
20	122	119	246	20	134	165	749
21	114	122	289	21	148	125	209
22	112	123	275	22	135	117	180
23	116	115	266	23	132	127	187
24	115	120	275	24	128	143	332
25	121	119	270	25	137	127	223
26	119	117	278	26	137	117	200
27	108	126	325	27	131	111	188
28	111	128	318	28	124	116	203
29	114	119	275	29	124	125	211
30	115	122	299	30	121	120	222
31	110	127	313	31	116	122	255
32	119	115	266	32	115	124	258

TABLE IV
VALUES OF TWO-BAND TEXTURE FEATURES

Uranium image no.	CON	IDM	ASM	Non-uranium image no.	CON	IDM	ASM
1	173	82	211	1	111	121	458
2	117	111	260	2	148	105	398
3	101	114	328	3	133	107	357
4	106	116	290	4	150	95	303
5	108	115	321	5	186	80	284
6	102	122	361	6	170	86	295
7	132	98	243	7	220	79	180
8	116	111	276	8	201	89	327
9	103	117	303	9	118	116	399
10	102	120	337	10	112	125	413
11	122	109	256	11	111	115	351
12	109	117	277	12	132	104	263
13	174	103	227	13	155	95	244
14	109	110	320	14	178	94	242
15	101	116	318	15	102	119	359
16	99	117	323	16	136	107	288
17	167	97	273	17	100	121	337
18	153	104	304	18	107	113	342
19	173	98	185	19	148	104	251
20	144	104	287	20	358	81	330
21	109	112	439	21	261	91	221
22	139	98	382	22	199	98	189
23	113	109	336	23	213	93	188
24	149	92	316	24	248	89	257
25	185	81	307	25	221	93	245
26	166	86	382	26	198	86	214
27	139	91	448	27	191	92	189
28	156	86	446	28	150	98	262
29	184	72	400	29	184	96	211
30	131	98	335	30	133	114	284
31	134	98	405	31	115	113	290
32	135	105	323	32	114	111	298

$\mathrm{CON} \equiv \Sigma\Sigma e_{ij}(i-j)^2$, and $\mathrm{ISM} \equiv \Sigma\Sigma e_{ij}/[1+(i-j)^2)$.] The features are based on pairs of pixels at unit horizontal separation. A good example of partial separation is given by the two-band CON feature, for which 11 of the 32 NU samples (numbers 5, 7, 8, 20–27), but none of the U samples, have values ≥ 186. None of the one-band features are as effective in defining distinctive NU ranges, though some of them do distinguish smaller (and different!) sets of NU's (e.g., for the IDM feature in eigenband 1, NU samples 1–3 and 8–11 have values ≥ 287, while no U samples are in this range). Another example, in which single-band features do better, is provided by the ASM feature in eigenband 3, for which 14 of the U samples (numbers 6, 10, 21–32) have values ≥ 265; two of the NU samples (20, 24) also lie in this range, but these NU samples have much higher values of the IDM eigenband 3 feature than any of the other samples. (The two-band ASM feature performs similarly, but not as well; it yields a distinctive range for only seven of the U samples: 22, 26–31.)

These results show that two-band features can sometimes yield better (partial) separations between texture classes than single-band features. They are thus potentially useful, and should be taken into consideration for other tasks involving texture classification on multispectral imagery.

IV. CONCLUDING REMARKS

Two-band texture features have the potential of providing textural information that is not available from single-band features. If other investigators experiment with two-band features, other cases will be found in which the two-band approach is advantageous.

ACKNOWLEDGMENT

The authors wish to thank D. Lloyd Chesley for help in preparing this paper.

REFERENCES

[1] R. M. Haralick, "Statistical and structural approaches to texture," *Proc. IEEE*, vol. 67, pp. 786–804, 1979.
[2] J. S. Weszka, C. R. Dyer, and A. Rosenfeld, "A comparative study of texture features for terrain classification," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-6, pp. 269–285, 1976.
[3] R. Chellappa, "Fitting random field models to images," Comput. Vision Lab., Comput. Sci. Center, Univ. of Maryland, College Park, TR-928, Aug. 1980.
[4] G. E. Lowitz, "Stability and dimensionality of Karhunen-Loeve multispectral image expansions," in *Proc. 3rd Int. Joint Conf. on Pattern Recognition*, 1976, pp. 673–677.
[5] *Geometric Pattern Recognition Techniques Applied to LANDSAT Digital Data for Uranium Exploration*, Earth Satellite Corp., Chevy Chase, MD, 1979.

Building and Road Extraction from Aerial Photographs

MOHAMAD TAVAKOLI AND AZRIEL ROSENFELD, FELLOW, IEEE

Abstract—A method of extracting features such as buildings and roads from high-resolution aerial photographs is described. The approach involves several successive stages of grouping of edge segments. Straight line