

Symbol discriminability models for improved flight displays

Albert J. Ahumada^a, Maite Trujillo San-Martin^b and Jennifer Gille^c

^aNASA Ames Research Center, MS 262-2, Moffett Field, CA 94035-1000

^bNational Research Council, NASA Ames Research Center, CA

^cUniversity of California, Santa Cruz, NASA Ames Research Center, CA

ABSTRACT

Aviation display system designers and evaluators need to know how discriminable displayed symbols will be over a wide range of conditions to assess the adequacy and effectiveness of flight display systems. If flight display symbols are to be safely recognized by pilots, it is necessary that they can be easily discriminated from each other. Sometimes psychophysical measurements can answer this question, but computational modeling may be required to assess the numerous conditions and even help design the empirical experiments that may be needed. Here we present an image discrimination model that includes position compensation. The model takes as input the luminance values for the pixels of two symbol images, the effective viewing distance, and gives as output the discriminability in just-noticeable-differences (d') and the x and y offset in pixels needed to minimize the discriminability. The model predictions are shown to be a useful upper bound for human symbol identification performance.

Keywords: Human vision models, cockpit display, symbology, image discrimination.

1. INTRODUCTION

Display technology advances afford the capacity to display traffic information in the cockpit using traffic symbols varying in shape and color to provide crucial information during flight operations. In this paper we present measurements of the ability of observers to identify such symbols and present a discriminability metric for the evaluation of the discriminability of symbol pairs. The goal of the metric is to provide objective measures of discriminability in support of design and manufacturing requirements for improved cockpit and air traffic management. The metric is based on our earlier work on image discrimination models for detection and discrimination¹⁻⁵ and some previous work on symbol discrimination.^{6,7}

Discriminability is only one component of what Yey and Chandra⁸ call distinctiveness, the degree to which the symbol can be identified by itself. Symbol identification involves many cognitive processes, such as feature learning, feature extraction, attention, and memory effects. Bruner⁹ gives a good overview of these higher-level processes that can affect symbol categorization. Here we are only trying to develop a simple symbol discriminability metric that only depends mainly on low-level visual processes. If such a metric predicts that symbols will be confused, they should indeed be confused, but if it says that they are discriminable, they may still be classified as the same symbol. We present a tool that could be used to measure the discriminability of pairs of stimuli. All pairs in a set of potential symbols would need to be compared to ensure discriminability, but discriminability would not ensure accurate categorization. For example, in the color domain, it is well known that many colors are discriminable from each other, but that relatively few color categories can be accurately recognized by naïve observers.^{10,11}

Our goal here is to extend an image discrimination model so that it can serve as a symbol discriminability metric. Image discrimination models take two images as input and predict their discriminability in just-noticeable-differences (JNDs) for a particular viewing condition. The luminance-only (achromatic), single-channel versions of these models have been successful at predicting the detectability of simple and complex

targets.¹⁻⁵ Watson has also extended them to predict letter identification.^{6, 7} Two modifications were needed. One modification in these applications is the inclusion of image translation. A letter can be recognized independent of its position in the image. The other modification is the addition of a response prediction mechanism. Watson added noise in the image domain and the response was selected by finding the best fitting response template. Although a response model may be necessary to simulate a pilot performance in an overall system evaluation, this part of the model is computationally intensive and not necessary if the goal is to only to ensure that all the symbols are discriminable from each other. The metric we propose is an image discrimination model that takes two symbols and computes the smallest JND between them over a range of symbol positions and symbol sizes. Because some of the symbols are rotated to indicate heading, the model should, in general, take rotation into account also. Here we have only compared the zero-heading-angle version of the symbols.

2. EXPERIMENTAL MEASUREMENT OF SYMBOL IDENTIFICATION

Two sets of symbol identification data will be presented. The first set was collected for the FAA by Zuschlag,¹² the second set is a partial replication of that study. The methods for the first study will be described first, followed by a description of the methodological differences in the replication.

2.1 Experimental Methods

2.1.1 Stimuli

The 19 symbol images appear in Figure 2.1.1. Their colors are described in Table 2.1.1.

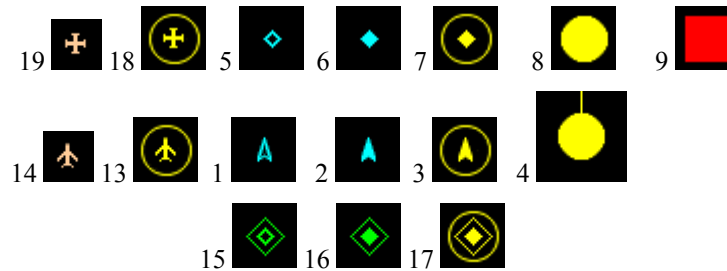


Figure 2.1.1: The Volpe experiment symbols.

Color	Symbol #	R	G	B
red	9	254	0	0
pink	14, 19	252	204	155
cyan	1, 2, 5, 6	0	254	252
green	10, 11, 15, 16	0	255	0
yellow	3, 4, 7, 8, 12, 13, 17, 18	255	255	0
black	Background	0	0	0

Table 2.1.1: Symbol set (color, stimuli index, RGB values for the color at its peak level)

Nine of the symbols (1-4, 10-14) are directional. The directional symbols were presented in one of six orientations on the flight display, but the orientation on the response display was always vertical.

2.2.2 Observers

Ten pilot observers were recruited at an airport. All had normal color vision and adequate visual acuity.

2.2.3 Procedure

On each trial, the observer was presented with one of the nineteen symbols in isolation on an Avidyne flight situation display (5RR-MFC-Series). The displayed image had a size of 4 in by 3 in, a resolution of 320 by 234 pixels, and thus a display resolution of 80 pixels per in. Four different viewing distances were used: 22, 44, 66, and 88 inches. At the 22 inch viewing distance the screen thus displayed 30.7 pixels per degree. The image was presented for 0.25 sec, preceded by a 1.5 sec fixation cross (horizontal and vertical white (255, 255, 255) 400 by 2 pixel lines) and followed by a crossed-hatch pattern. The observer used a mouse to respond by clicking on one of the 19 symbols presented continuously on a separate laptop display i.e.: Panasonic Toughbook (CF-37). The symbols were displayed on the laptop in the arrangement of Figure 2.1.1. Each observer responded to six replications of each stimulus, (6 x 19 x 4 = 456 trials). The oriented symbols were presented on the flight display in one of six orientations (45, 90, 135, 180, 225, or 315 degrees) selected at random, but the orientation on the response display was always vertical. Table 2.2.3 shows the sequence of the 4 distance conditions for each of the 10 observers.

Distance\Observer (in)	1	2	3	4	5	6	7	8	9	10
22	1	1	3	4	4	1	2	3	3	4
44	2	2	4	1	3	4	1	2	1	2
66	3	3	1	2	2	3	4	1	4	1
88	4	4	2	3	1	2	3	4	2	3

Table 2.2.3: Sequence of distance conditions for each observer.

2.2.4 Replication method variations

One observer responded to 10 repetitions of each symbol on 7 different days at a distance of 110 inches, which gave the same pixel resolution in pixels per degree as the 88 inch distance in the original experiment. The primary display was a Sony Trinitron (GDM-FW900, RGB line pitch = 0.23 mm). The display controller resolution was set to 1024 by 768 pixels and the pixel size was set to 64 pixels per inch in both directions. The observer's mouse controlled a cursor on another Sony Trinitron (CPD-200SX, Dot pitch = 0.25 mm). The symbol exposure duration was 0.5 sec, instead of 0.25 sec. Symbols were chosen randomly with replacement from the 190 stimuli presented on each day.

2.3 Results

Table 2.3.1 shows the number of valid trials per observer for each of the distances in the original experiment.

Distance\Observer (in)	1	2	3	4	5	6	7	8	9	10
22	114	114	114	114	114	114	113	114	114	114
44	114	114	114	114	113	114	114	114	114	106
66	114	114	114	114	114	114	114	114	113	114
88	114	114	112	114	113	114	114	114	114	112

Table 2.3.1: Number of valid trials per observer for each distance.

Table 2.3.2 shows the number of errors made by each observer in each of the distance conditions. The last line shows the number of color errors at the 88 in distance.

Distance\Observer		1	2	3	4	5	6	7	8	9	10
(in)											
22		4	1	6	11	7	1	3	0	0	3
44		4	3	3	21	10	2	8	0	0	2
66		19	11	28	31	58	45	4	25	6	27
88		42	33	31	56	86	70	46	47	33	51
Col	88	8	1	0	6	36	8	7	4	2	6

Table 2.3.2: Number of errors made by each observer at each distance

In this data set, only the 88 in distance consistently resulted in errors for all observers. The 66 inch distance errors correlated strongly with the temporal position of that condition in the experiment ($n=10$, Spearman's $\rho = -0.56$). Observer 5 was the only one to have the 88 in distance condition first and made a much higher number of symbol and color confusions in that condition than any other observer. We now restrict our attention to the confusion matrix for the 88 in distance and the best 9 observers (Table 2.3.3). Rows represent symbols presented; columns represent responses made.

	Red			Cyan				Green				Yellow							
	9	14	19	1	2	5	6	10	11	15	16	4	8	12	17	18	7	13	3
9	59	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14	0	39	9	2	0	3	0	1	0	0	0	0	0	0	0	0	2	1	0
19	0	31	19	0	0	4	1	0	0	0	0	0	0	0	0	1	0	0	0
1	0	0	0	22	13	1	0	6	3	1	2	0	0	0	0	0	0	0	0
2	0	0	0	5	35	0	5	2	8	0	1	0	0	0	0	0	0	0	0
5	0	0	0	2	0	31	10	0	0	4	6	0	0	0	0	0	0	0	0
6	0	0	0	0	8	2	30	1	0	0	4	0	0	0	0	0	0	0	0
10	0	0	0	0	1	0	0	7	39	0	0	0	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0	6	41	1	0	0	0	0	0	0	0	0	0
15	0	0	0	0	0	0	1	1	3	13	39	0	0	0	0	0	0	0	0
16	0	0	0	0	0	0	0	2	3	6	35	0	0	1	1	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	54	2	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0	0	0	0	1	46	0	0	0	0	0	0
12	0	0	0	0	0	0	0	1	0	0	0	0	0	38	0	1	0	2	11
17	0	0	0	0	0	0	0	0	0	0	0	0	0	10	29	1	6	11	3
18	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	14	5	33	4
7	0	0	0	0	0	0	0	0	0	0	0	0	0	2	2	28	13	12	
13	0	1	0	0	0	0	0	0	0	1	0	0	4	2	2	12	34	13	
3	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	1	4	39

Table 2.3.3: Confusion matrix for the 88 in distance and the best 9 observers. Each row has, for a symbol presented, the total number of times each of the 19 possible responses was given.

The information transmitted^{10, 11, 13} in this confusion matrix is 2.2 bits per symbol, the information transmission rate for 7 errorless symbols ($\log_2 7$). The information in a discrete probability or frequency distribution with probabilities p_i , $i = 1, n$, is given by Eq. 2.3.1.

$$H(p) = -\sum_i \log_2 p_i \quad (2.3.1)$$

Associated with a confusion table are three distributions, the proportion of times each symbol was presented, the proportion of times each response was made, and the proportion of times each symbol presentation and response pair occurred. If we call the information in these distributions $H(s)$, $H(r)$, and $H(s,r)$, then the information transmitted in the confusion matrix is given by Eq. 2.3.2.

$$T(s, r) = H(s) + H(r) - H(s, r) \quad (2.3.2)$$

The first row and column show that the red square (9) was not confused with any other symbol. The large yellow circle (8) was confused 3 times with the same circle with the orientation line (4), but only rarely and the two were never confused with any other symbol. This 2 by 2 confusion matrix can be converted to a d' value (Eq. 2.3.3).

$$d' = z(\text{Pr}(8|8)) - z(\text{Pr}(8|4)) = z(46/47) - z(2/56) = 2.03 - (-1.80) = 3.83 \quad (2.3.3)$$

Although this 2 by 2 matrix is roughly symmetric, many of the confusions were not. The only two pink symbols, the plane (14) and the cross (19) were usually confused with each other, but both were more frequently reported to be the plane symbol. The same tendency was shown when they were changed to yellow and surrounded by a circle (13 and 18). Another consistent response bias is that when an open figure and its filled version are confused, the response tends to be ‘filled’. This is illustrated by the confusions between the diamonds (5 and 6), the triangles (1 and 2), the outlined diamonds (15 and 16), and the outlined triangles (10 and 11). The confusions that suggest the need for size scaling are those in which a shape was given the response of its outlined version. Even though the color of the outlined versions is green rather than cyan, the triangles and diamonds alone were sometimes responded to as their outlined version.

Table 2.3.4 shows the confusion matrix for the second data set. For this matrix the information transmitted per symbol is 3.7 bits, corresponding to 13 errorless symbols. The data for the second set are also organized by color, and in this data set there were no confusions between different colors.

	Red			Cyan				Green				Yellow							
	9	14	19	1	2	5	6	10	11	15	16	4	8	12	17	18	7	13	3
9	70	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14	0	69	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
19	0	18	52	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0	0	64	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	1	69	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	5	0	57	8	0	0	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0	11	0	59	0	0	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	52	18	0	0	0	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0	13	57	0	0	0	0	0	0	0	0	0	0
15	0	0	0	0	0	0	0	0	0	45	25	0	0	0	0	0	0	0	0
16	0	0	0	0	0	0	0	0	0	9	61	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	69	1	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0	0	0	0	0	70	0	0	0	0	0	0
12	0	0	0	0	0	0	0	0	0	0	0	0	0	70	0	0	0	0	0
17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	70	0	0	0	0
18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	45	19	3	2
7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	66	0	3
13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	8	19	42
3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	68

Table 2.3.4: Replication Data Confusion matrix

3. DISCRIMINATION MODELS

Here we describe in detail the computations performed by the modeling program and the parameter values. The model is available at the first author’s NASA website¹⁴. The model program takes as input a pair of images. Both images are converted from color to luminance. Next they are converted to visible contrast images, taking into account the contrast sensitivity of the visual system and contrast-gain masking. The translation that minimizes the Euclidean distance between the visible contrast images is then found, and that minimum distance is scaled into a number of JNDs (d') by the assumption that 1 JND is $10^{-6} \text{ deg}^2 \text{ sec}$ of contrast energy (0 dB).^{3,5}

3.1 Color to luminance conversion

The conversion from color RGB values (0-255) to luminance (cd/m^2) was done using the simple color gamma function shown in Eq. 3.1.1.

$$L = L_0 + L_R (R/255)^\gamma + L_G (G/255)^\gamma + L_B (R/255)^\gamma \quad (3.1.1)$$

Table 3.1.1 gives the values of these parameters for the two primary displays.

Display	L_0	L_R	L_G	L_B	γ
Avidyne	7.1	26.7	85.9	14.0	2.71
Sony	2.8	15.6	54.0	6.8	2.53

Table 3.1.1: Gamma function parameters for the two primary displays.

After conversion to luminance, the images were extended from their original sizes of 25 by 25 or 32 by 32 pixels to 128 by 128 pixels by imbedding them in a 128 by 128 matrix of L_0 values. This was done to minimize wraparound effects from the FFT-based filtering. Next the images were pixel-replicated by a factor of 4 so that the position optimization would occur at quarter-pixel resolution.

3.2 Visible contrast calculation

The luminance image was first blurred by a Gaussian with a standard deviation of 1.414/60 deg. Next a background image B was computed by mixing the image L with a proportion $p_L = 0.8$ of the background,

$$B = p_L L + (1-p_L) L_0 \quad (3.2.1)$$

Next the background image is blurred by a Gaussian with a standard deviation of 16 times 1.414/60 deg. Then a contrast image C (Eq. 3.2.2) is formed by point-by-point division of these two images.

$$C = L / B - 1 \quad (3.2.2)$$

A contrast-energy image E is then formed by point-by-point squaring of C, as shown in Eq. 3.2.3.

$$E = C^2 \quad (3.2.3)$$

This image is then blurred with the same Gaussian used to blur B. And finally the blurred contrast image E is used to adjust the contrast by the contrast gain function applied point-by-point, giving the visible contrast image V (Eq. 3.2.4).

$$V = C / (1 + k E)^{0.5} \quad (3.2.4)$$

The value of k was set to 400 corresponding to a masking RMS threshold contrast of 5%.

3.3 JND calculation

The minimum distance between two images can be found quickly by finding the maximum of the cross correlation between the two images in the Fourier domain. The minimum distance between two visible contrast images is then converted to JNDs in d' units as follows (Eq. 3.3.1).

$$d' = S |V_1 - V_2| \quad (3.3.1)$$

Where the contrast sensitivity $S = (t / 10^{-6} \text{ deg}^2 \text{ sec})^{0.5}$. The variable t is the duration of the symbols, 0.25 sec in the original experiment and 0.5 sec in the replication, giving values of S of 500 and 707, respectively.

3.4 Model predictions compared with results

Since the current model does not include color differences, the observer data to be predicted will be limited to the confusions among pairs of the same color. There are only a few same-color pairs in which the responses to the pair are only limited to the pair, allowing a d' value to be computed using the formula above. For all the pairs, it is possible to compute the information transmitted from that pair of symbols to all responses. Table 3.4.1 shows the information transmitted for all the same color pairs for the original experiment. Table 3.4.2 shows the same table for the replication.

pink							
19							
14 0.10							
cyan							
2 5 6							
1	0.25		0.74		0.65		
2			0.78		0.43		
5					0.45		
green							
11 15 16							
10	0.22		0.81		0.76		
11			0.76		0.72		
15					0.05		
yellow							
8 12 17 18 7 13 3							
4	0.81	1.00	1.00	1.00	1.00	1.00	1.00
8		1.00	1.00	1.00	1.00	1.00	1.00
12			0.51	0.68	0.69	0.52	0.38
17				0.41	0.39	0.30	0.58
18					0.27	0.13	0.57
7						0.14	0.43
13							0.35

Table 3.4.1: Information transmitted in bits per symbol for all same-color pairs in the original experiment.

pink							
19							
14 0.49							
cyan							
2 5 6							
1	0.73		0.82		0.89		
2			0.97		0.67		
5					0.75		
green							
11 15 16							
10	0.24		1.00		1.00		
11			1.00		1.00		
15					0.22		
yellow							
8 12 17 18 7 13 3							
4	0.95	1.00	1.00	1.00	1.00	1.00	1.00
8		1.00	1.00	1.00	1.00	1.00	1.00
12			1.00	1.00	1.00	1.00	1.00
17				0.95	1.00	1.00	1.00
18					0.45	0.61	0.87
7						0.61	0.87
13							0.18

Table 3.4.2: Information transmitted in bits per symbol for all same-color pairs in the replication.

Table 3.4.3 shows the JNDs (d' values) for the same-color pairs based on the calibration of the Avidyne display. The calibration values were so similar for the Sony display used in the replication that the correlation between the two sets of predictions was 0.9997. The model predictions for all 13 pairs of the not circled symbols of the same color (1 pink, 6 cyan, 6 green) and the 15 yellow pairs of circled symbols (except the two solid circles) are shown in Figure 3.4.1 together with the observer performance on the same pairs.

4. DISCUSSION AND FUTURE WORK

The model appears to perform as well as it could. In principle, performance can not be better than image discrimination performance, feature extraction and memory effects should only make performance worse. It is surprising that the performance actually reaches the discrimination limit. The model discrimination threshold is as low as has ever been reported for observers. The main feature of the model that could predict worse performance than that of the observers is the wide spread of contrast energy masking. Perhaps there is also some capitalization on the effects of random error in the information transmission scores. The model was run with size compensation as well as position compensation with no improvement in the predictions. However, the only examples of confusions that would require the size compensation are cross-color confusions. The size compensation will be tested again when we complete the color discrimination version of the model. There are two other features of the experiment that are not incorporated into the model.^{15, 16} The luminance and contrast energy masking by the fixation cross hair was ignored, as was the fact that the oriented symbols were presented in different orientations.

The model can be used directly to flag pairs for which the discriminability is poor, such as the two pink symbols 14 and 19. In conjunction with the behavioral assessment, it can be used to flag pairs where the discriminability is good, but the performance is poor, such as the confusions of 12 with 3, 13 and 17. These look like pairs for which the size compensation might help, but that the real problem is in feature extraction.

ACKNOWLEDGEMENTS

Thanks are due to William Krebs and Bill Kaliardos for providing the rational and the FAA support for this project, to Andrew Watson for showing that the discrimination models with translation invariance could predict letter confusions, to Michael Zuschlag for providing the raw behavioral and calibration data from his experiment, and to Bettina Beard for providing the laboratory facilities at NASA Ames.

REFERENCES

1. Rohaly, A. M., Ahumada, A. J. Jr., & Watson, A. B., "Object detection in natural backgrounds predicted by discrimination performance and models", *Vision Research*, vol. 37, 3225-3235, 1997.
2. Ahumada, A. J. Jr., Beard, B. L., "A simple vision model for inhomogeneous image quality assessment", in *SID Digest of Technical Papers*, ed. J. Morreale, vol. 29, Paper 40.1, Santa Ana, CA, 1998.
3. Watson, A. B., Visual detection of spatial contrast patterns: Evaluation of five simple models, *Optics Express* 6(1), 12-33, 2000.
4. Beard, B. L., Jones, K. M., Chacon, C., and Ahumada, A. J., Jr., "Detection of blurred cracks: A step towards an empirical vision standard", Final Report for FAA Agreement DTFA-2045, 2005.
5. Watson, A. B., Ahumada, A. J. Jr., A standard model for foveal detection of spatial contrast, *Journal of Vision*, 5 (9), 717-740, 2005.
6. Watson, A. B., and Ahumada, A. J., "Human optical image quality and the Spatial Standard Observer", OSA Fall Vision Meeting, 2004.
7. Watson, A. B., and Ahumada, A. J., "Predicting acuity from aberrations with the spatial standard observer", *Investigative Ophthalmology and Visual Science* 46, E-Abstract 3614, 2005.
8. Yeh, M., and Chandra, D., "Issues in symbol design for electronic displays of navigation information", *Proceedings of the 23rd DASC Conference*, 24-28 October, Salt Lake City, UT, 2004.
9. Bruner, J. S., "Beyond the information given: Studies in the psychology of knowing", Oxford, UK: W. W. Norton, 1973.

10. Miller, G., "The magical number seven, plus or minus two". Psychol. Rev. 63, 81-97, 1956.
11. Garner, W. R., "Uncertainty and Structure as Psychological Concepts", New York: John Wiley and Sons, Inc, 1962.
12. Zuschlag, M., "Assessment of proposed traffic symbol set". FAA Human Factors Research and Engineering Division FY2004 Report, p. 44. <https://www.hf.faa.gov/docs/508/docs/2004Report.pdf>, 2004.
13. Shannon, C. E., Weaver, W., The mathematical theory of communication, Urbana: University of Illinois Press, 1949.
14. Ahumada, A. <http://vision.arc.nasa.gov/~al/code/>
15. Ahumada, A. J. Jr., and Krebs, W. K., "Masking in color images", in B. E. Rogowitz and T. N. Pappas, Eds., Human Vision and Electronic Imaging VI, SPIE Proc. Vol. 4299, pp. 187-194, 2001.
16. Wuerger, S., Watson, A. B., and Ahumada, A. J. Jr. "Toward a standard observer for spatio-chromatic detection", SPIE Proc. Vol. 4662, 159-172, 2002.