26.541
VSS 2011

# Characterizing the salience and interactions of informative image statistics


Jonathan D. Victor, Daniel J. Thengone, Mary M. Conte

Department of Neurology and Neuroscience, Weill Cornell Medical College


## Motivation and Overview

**Natural scenes have complex statistical structure, including correlations of low and high order. However, most models of visual processing are based on stimuli with simple statistical structure, such as gratings and white noise.**

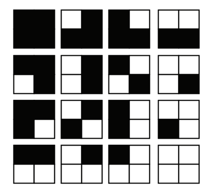
- Since the complex statistics inherent in natural scenes are absent from the stimuli used to construct most models, it is not surprising that the models fail to fully account for responses to natural scenes.
- This motivated us to develop a library of stimuli that abstract the statistics in natural scenes, to enable testing their effects in isolation, and how they interact.
- To reduce the dimensionality of the problem, we focus on the local statistics of binarized images. This leads to a space of visual textures; the 10 coordinates of this space completely capture correlations of orders 1 through 4.

**Extraction of local orientation signals is key to early visual processing.**

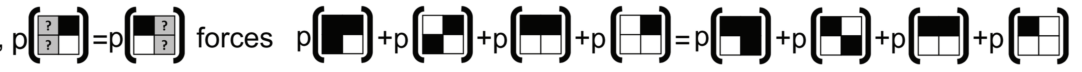
- In natural images and in the texture space, tokens at multiple orientations are simultaneously present, and can overlap in space. Moreover, orientation information is carried not only by second-order statistics, but also by higher-order ones.
- We therefore examine how multiple orientation signals interact, and whether this interaction depends on the kind of orientation information.
- Second-order statistics are generally processed in a way that keeps each orientation separate.
- Third-order statistics are processed in a way that pools across orientations.

## Coordinates for a Local Binary Texture Space

Textures are defined by the distribution of colorings of 2x2 blocks. There are 16 (=$2^{2x2}$) possible colorings (right). However, there are constraints among these probabilities because the probabilities of smaller blocks cannot depend on *where* they occur within the 2x2 block.



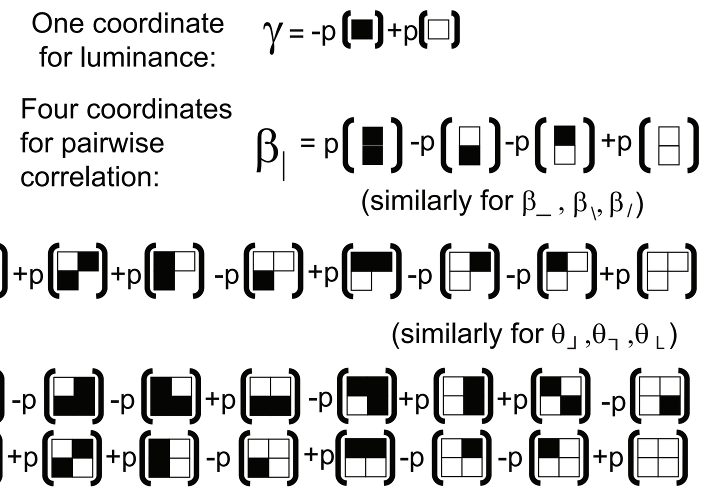
For example,



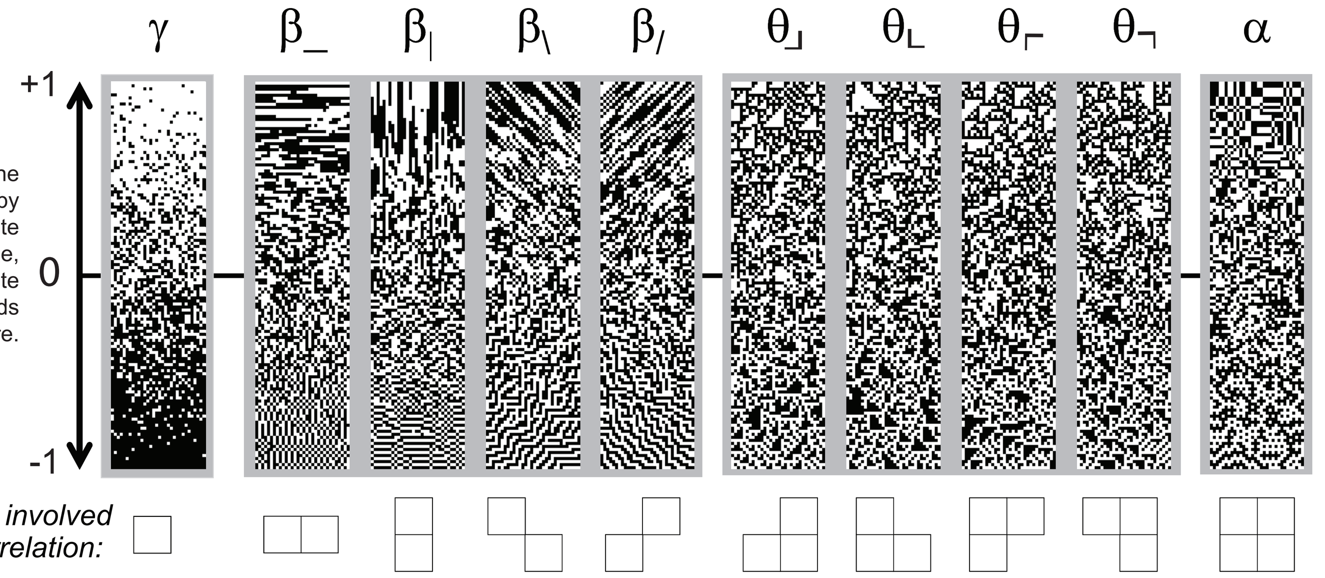

There are 6 independent linear constraints on 16 probabilities, yielding 16–6 = 10 free parameters.

### What are the 10 coordinates?

The 10 independent parameters are captured by 10 coordinates. Each is a linear combination of the probabilities. Each coordinate ranges from -1 to 1, and has a value of 0 for the random texture.




### Coordinate Axes




## Methods and Psychometric Functions

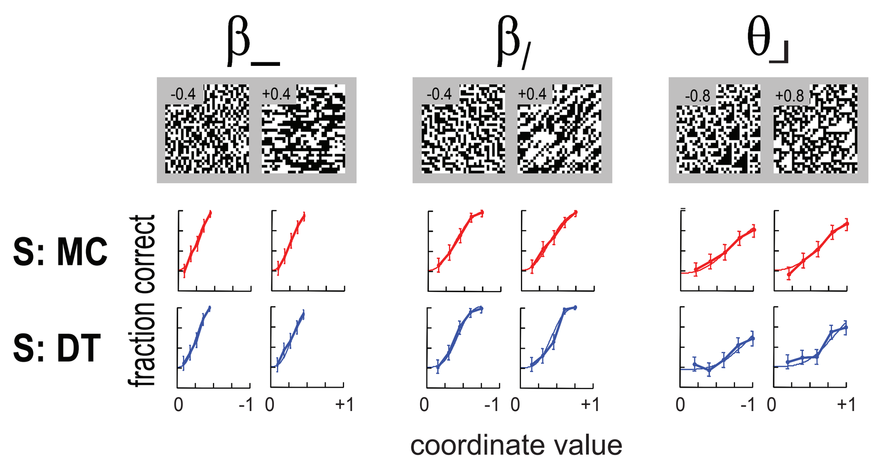
**SUBJECTS**
12 subjects
VA: 20/20, with correction if needed
Practice: approx 1600 trials

**CONDITIONS**
8 repeats of 20 on-axis points
16 repeats of 8 off-axis points
288 trials per block, random order
15 blocks = 4320 trials per plane
Feedback during practice only

**STIMULI**
Pixel Size: 14 min
Display Size: 14.8 $deg^2$
Binocular viewing at 1m
Contrast: 1.0
Duration: 120 ms (followed by mask)
Target: 16 x 64 pixels on a 64 x 64 array
Trials either have a structured target on a random background, or random target on structured background

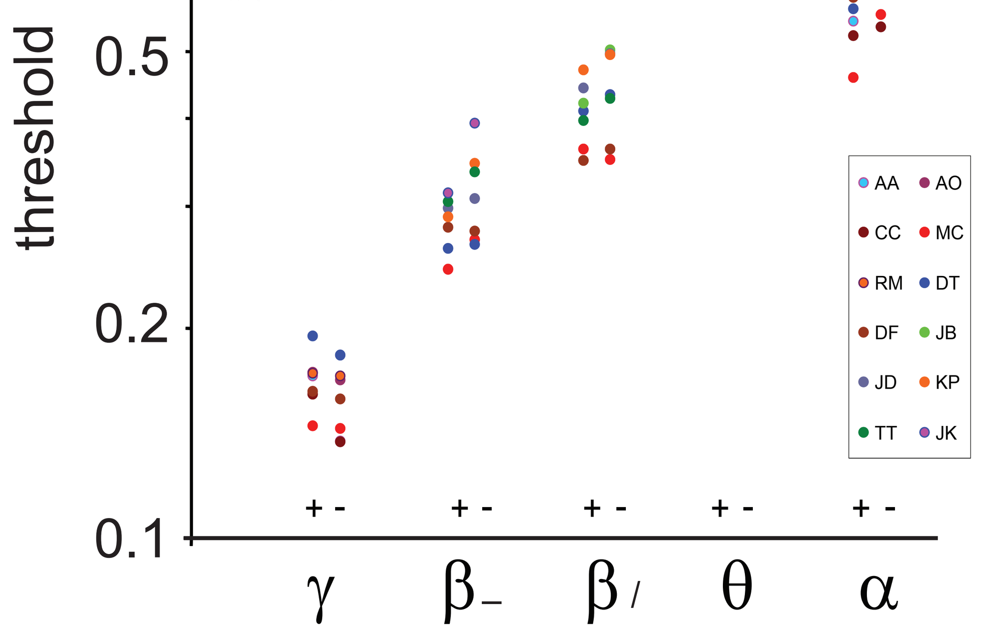
**TASK**
Find the location of the target stripe (4-AFC, top, right, bottom, left)

In this sample stimulus, the target stripe (defined by $\theta_\lrcorner$) is on the bottom.




Each plot shows the fraction correct for stimuli that vary along a single texture coordinate. Performance is similar for positive and negative excursions of a coordinate, and was highly consistent across subjects (shown: MC and DT). Curves are maximum-likelihood fits to Weibull functions (shape parameter typically 2.2 to 2.6). Error bars are 95% confidence limits.

Thresholds are highly consistent across subjects and are similar for positive and negative excursions, in approximate ratio 1:2:5:4 ($\gamma$:$\beta$:$\theta$:$\alpha$).



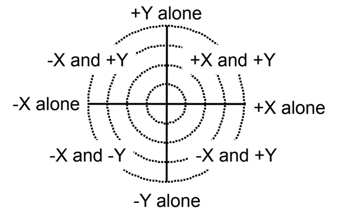

## Analyzing How Image Statistics Interact

To analyze how image statistics interact, we determine their ability to drive segmentation, alone and in graded pairwise combinations.

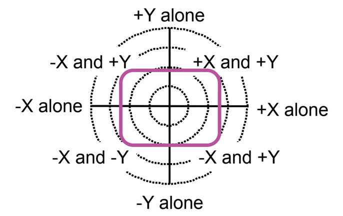
Results are summarized by a polar plot of the isodiscrimination contour

- Distance from the origin represents the threshold
- Direction from the origin represents the ratio of the two image statistics
- X- and Y- axes correspond to thresholds for each image statistic alone
- Oblique directions correspond to thresholds for the image statistics in combination

The shape of the isodiscrimination contour indicates how the image statistics interact.

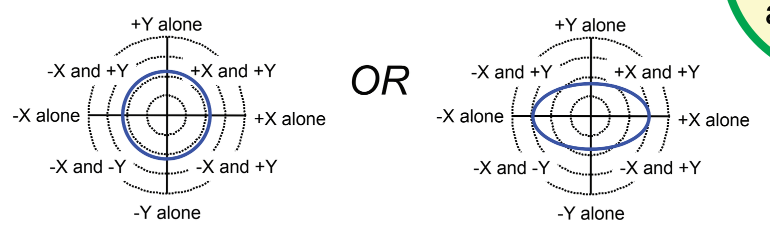



One way that image statistics might interact is via "winner-take-all": only the largest signal matters. In this case, there is no subthreshold summation. Since threshold is reached only when one of the statistics reaches threshold by itself, the resulting isodiscrimination contour will be box-like.

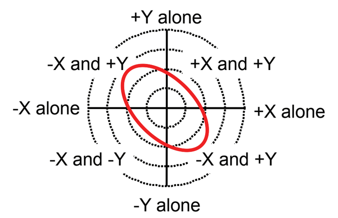


***This is not what we find.***

If subthreshold summation is present, contours will be more elliptical. Circular elliptical contours correspond to a Euclidean perceptual metric, with salience equally determined by amount of each statistic present: $(X^2+Y^2)^{1/2}$. If one image statistic is more salient than the other, the contour is elliptical, oriented along the coordinate axes.




This is what we find for second-order statistics $\beta_{\_}\beta_|$ and $\beta_{\_}\beta_/$ .

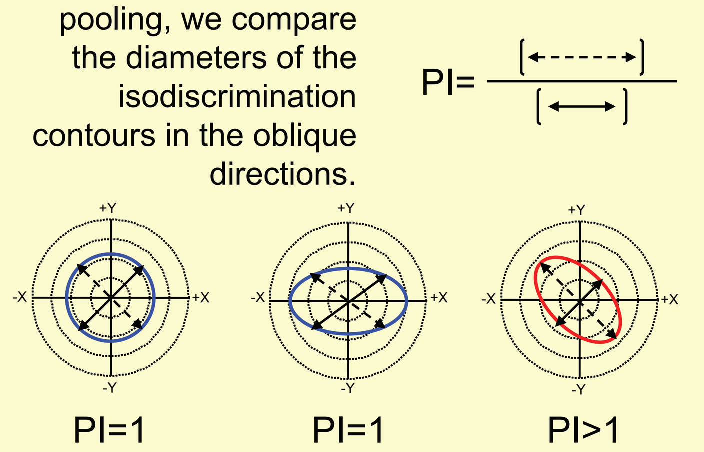


This is what we find for third-order statistics $\theta_\lrcorner\theta_\llcorner$ and $\theta_\lrcorner\theta_\ulcorner$ .

Another possibility is that X and Y are pooled. This could happen, for example, if they are detected by the *same* mechanism. In this case, thresholds are reduced when X and Y have the same sign (because signals add up), and are increased when they have opposite sign (because signals cancel). This results in a tilt of the isodiscrimination contour, with elongation into the quadrants in which X and Y have opposite sign.

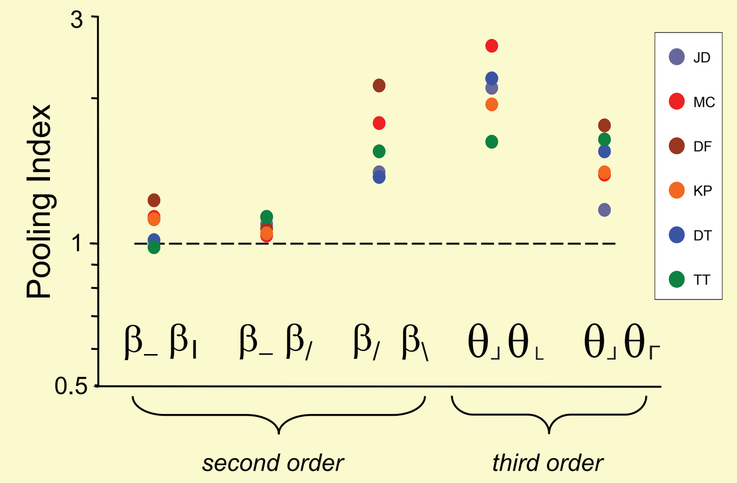
## Pooling Index

To quantify the extent of pooling, we compare the diameters of the isodiscrimination contours in the oblique directions.




This ratio, the "pooling index", is 1 if image statistics are processed independently, and >1 if they are pooled.

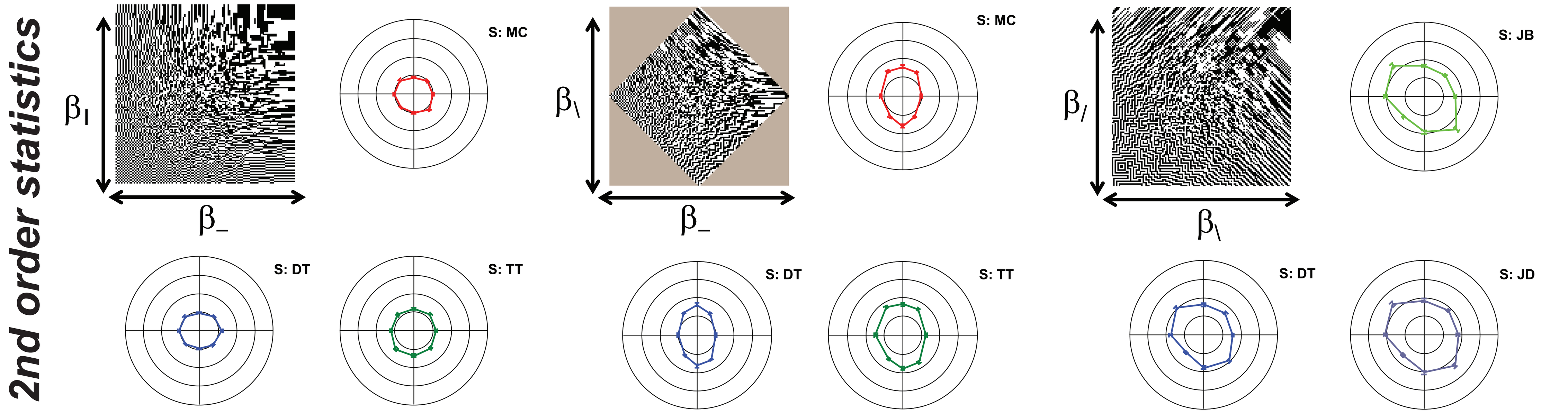
## Pooling Index Summary




The pooling index shows which pairs of image statistics are processed separately (PI=1), and which are pooled (PI> 1). Second-order statistic pairs $\beta_{\_}\beta_|$ and $\beta_{\_}\beta_/$ are processed separately; third-order statistic pairs $\theta_\lrcorner\theta_\llcorner$ and $\theta_\lrcorner\theta_\ulcorner$ and second-order statistic pair $\beta_/\beta_\backslash$ are pooled.

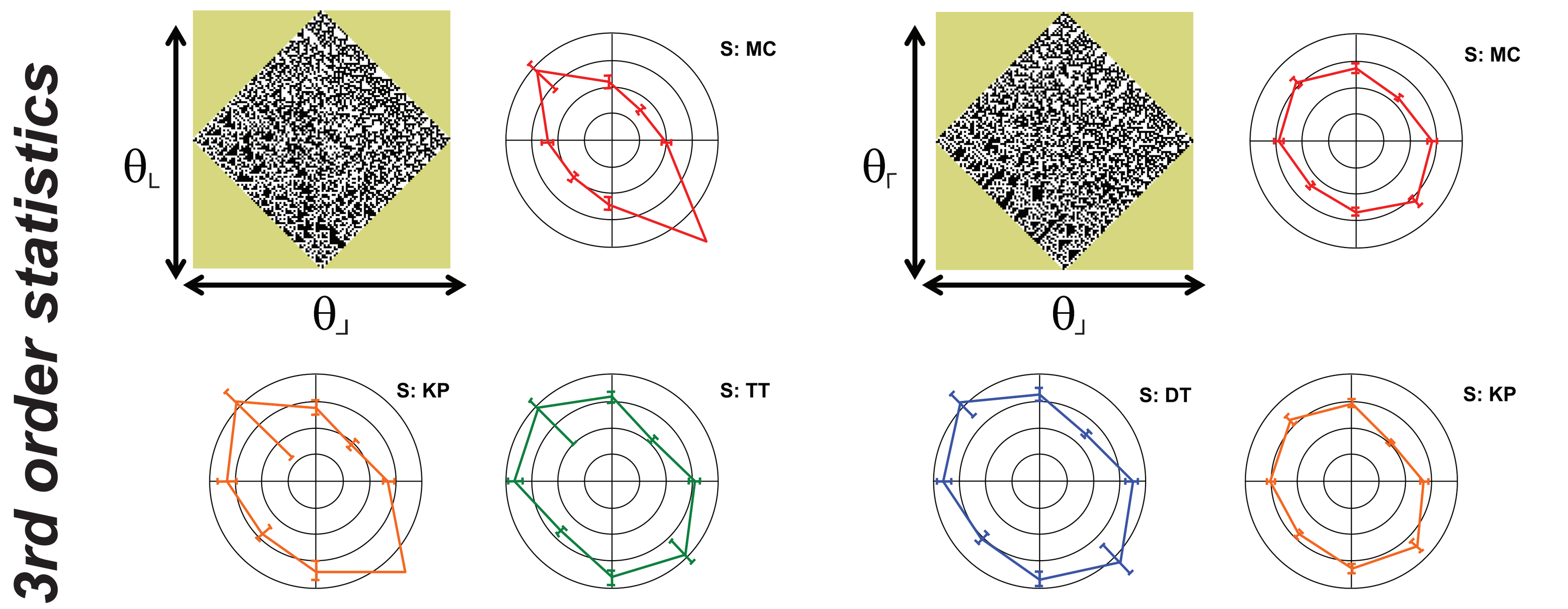
## Summary and Conclusions

- To study how image statistics interact, we abstract the complex statistical structure of natural scenes into a reduced space of binary textures with local correlations.
- Within this 10-dimensional space, sensitivities to individual image statistics are highly consistent within and across N=12 observers. Thresholds for first-, second-, third-, and fourth- order statistics are in the ratio of 1:2:5:4.
- Tokens representing multiple orientations can coexist, and orientation information can be carried by second- and higher-order statistics.
- The perceptual interactions of these image statistics are also highly consistent across N=6 observers, and correspond to an approximately Euclidean perceptual space.
- Orientation signals carried by second-order statistics are generally processed independently, while orientation signals carried by third-order statistics are pooled.




### Coordinate Planes
### simple interactions between statistics

Each panel illustrates a texture coordinate pair that was studied. Within each panel, the texture image indicates the gamut of the coordinate pair: the center of each image is random; the border indicates the limits imposed by consistency conditions. One can get an idea of the salience of each coordinate and how they interact by noting where each image becomes visibly non-random. This is quantified by isodiscrimination contours, shown for three subjects adjacent to the texture image.





Supported by EY7977