ASIAN JOURNAL OF PHARMACEUTICAL AND CLINICAL RESEARCH

INNOVARE ACADEMIC SCIENCES
Knowledge to Innovation

Online - 2455-3891
Print - 0974-2441

Special Issue (April)

Full Proceeding Paper

SHAPE SELECTIVITY OF AND SYNTHETIC VISUAL INTELLIGENCE

ANKUSH RAI*, JAGADEESH KANNAN R

School of Computing Science and Engineering, VIT University, Chennai, Tamil Nadu, India. Email: ankushressci@gmail.com

Received: 13 December 2016, Revised and Accepted: 03 April 2017

ABSTRACT

The process of coding information for face recognition in human is largely remaining unknown. In this study, we carry out few experiments to determine the factors influencing coding mechanism in parahippocampal place area of the brain. The results show some significant outcome toward the shape selectivity of the brain and latter we construct a computational mechanism to mimic the coding features behind face recognition.

Keywords: Computational neuroscience, Face recognition, Parahippocampal place area.

 DOI: http://dx.doi.org/10.22159/ajpcr.2017.v10s1.19656

INTRODUCTION

When it comes to facial recognition, a contrast negated image is equivalent in preserving the information as it with the positive counterpart. An image's 2D geometric and spectral structure are completely preserved in negation. However, this has noticeably reverse consequences on human's ability to recognize faces [1-6]. Thereby investigating the reason behind the phenomenon is essential for developing an understanding of the visual system applicable for face identification. Evidently, researchers have hypothesized for the phenomenon that the unnatural shading cues compromises shapes from shading processes [7-11]. However, it is still unclear whether this explanation sufficed. Particularly, in the manifestation of experimental results where recognition performance in an unavailability of the shading gradients [12], and models of face recognition that are based on the use of 2D intensity patterns than that of recovered 3D geometries [13,14]. In addition, negation forms faces to have odd pigmentation [15,16]; which has little effect on recognition performance [11]. It is vague that whether the viewer can beneficially extract information of pigmentation pattern formed across varied illumination states [17,18]. Of the infinitely many aspects of facial photometry, this study reveals that destruction of small sets of 2D geometric shapes adversely affects face-recognition performance as we have found that the brain bears a relationship between shape-selective recognition of objects in parahippocampal place area (PPA) that underlie the negation-induced decrements.

STIMULI AND PROCEDURE

Subjects

For each of the experiment involving different patterns of emerging of the images, human subjects were selected from a random pool of 20 subjects (10 females) belonging to 20-35 years of age. Note that, it is taken care of the fact that every subject had normal visual acuity and zero history of neuropsychological disorder of any sorts.

Experiment

Two stimulus sets of 20 images (of 800×600 pixels; with a display resolution of 38.4 pixels per degree), mainly including multiple natural objects of either rectilinear or round shape were showcased within a circular aperture of diameter >25° to the subjects in the experiment. Furthermore, the subjects using machine vision algorithm to track the position of the subjects gaze during the experiment are also recorded; such as to extract the correlation between the subjects' gaze and its functional magnetic resonance imaging (fMRI) data. Latter, the fMRI data of the PPA and the tracked location of the subject's gaze is linked to the computational unit; such that the artificial neural network was programmed in a way to get trained in real time by focusing on two factors: (1) The cross-section and strength of activation of the PPA regions with respect to time (stored in I matrix) (2) Subject's gaze localization with respect to time (stored in M matrix). The location of illuminant source varies from five places while maintaining the difference between traverse origin of the image on the screen; based on that the experimental reading is divided into 1-5 sets. Finally, twelve runs of the read ing are recorded (duration 3 s/image) in each experiment. The summary for specific stimuli in detecting specific geometrical shape is shown in Table 1.

ARTIFICIAL NEURAL SYSTEM: DATA GATHERING AND REPLICATION BY TRAINING OF REAL-TIME DATA

The artificial neural rule-based system used in the experiment employs the concept of neural network as well as fuzzy logic. The neural network used in work contains an input layer and an output layer. The number of neurons in hidden layers decides the objects to be classified. Output layer of the network is used to replicate the recognition process by carrying out the same task as the similar data to the subject is fed into the system. At time t each of the two elements of the focusing factors is processed within the layers (using matrix representation) that are related to the classes of rectilinear and curvilinear objects, and the output is manifested in the form of step-by-step identification process. The fuzzy learning mechanism is used between the weights of the input and middle layer to detect how often the outputs win the competition. Multilayer feed forward neural network is used in the first step during the examination. The input layer of neural network has M number of neurons, and the hidden layer has N_g neurons [19-21]. The output layer of the network has N neurons. Training of artificial neural network is done using Back propagation algorithm as modeled below:

Step 1: Develop a network with suitable number of neurons and other parameters as per value of I and M supplied.

Step 2: Analyze input image and map all detected and segmented objects and numbers into linear arrays.

Step 3: Read desired output converting each segmented object to a binary Unicode value. Characters are individually stored.

Step 4: Generate arbitrary weights within the interval [0, 1] and assign to all neurons in hidden layers and also output layer. Maintain a unity value weight for all neurons of the input layer.

Step 5: For each segmented object:

i. The output of the feed forward network is calculated.

ii. A comparison is made with the desired output corresponding to the symbol and compute error.

iii. Errors are propagated back across each neuron in previous layers

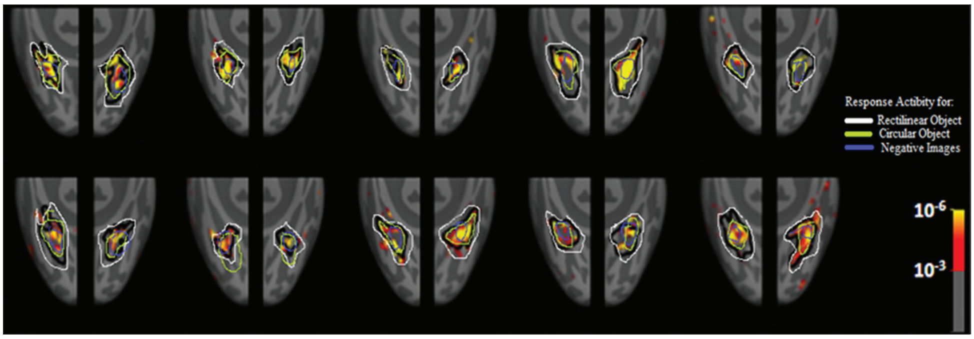

Fig. 1: Stimuli results representing activation area of shape-selective areas in parahippocampal place area for images with rectilinear, curvilinear and negative images where corresponding contour lines represent group average contour maps produced by recognition process of similar sets of images

to adjust the weights.

Step 6: The training dataset I is fed to the classifier and determine back propagation (BP) error by:

$$BP_{err} = C_{tar} - C_{out} \tag{1}$$

Where, C_{tar} is the desired target output and C_{out} is the actual network output. The value of C_{out} is determined as:

$C_{out} = \left[Y_2^{(1)} Y_2^{(2)} \dots Y_2^{(N)}\right]$ where $Y_2^{(1)}, Y_2^{(2)}, \dots, Y_2^{(N)}$ are the network outputs of each neuron. The individual network outputs can be computed as:

$$Y_2^{(1)} = \sum_{r=1}^{N_g} w_{2r1} Y_1(r) \tag{2}$$

$$Y_1(r) = \frac{1}{1+exp(-w_{1r1}.C_{in})} \tag{3}$$

Where, w_{2r1} is the weight of the connection from the *2r*th input element to the *1*th hidden unit. Equation 5.18 and equation 5.19 are activation functions of output layer and hidden layer, respectively.

Step 7: Adjust the weights of all neurons by w=w=Δw, where Δw is the change in weight estimated as:

$\Delta w = \gamma.\gamma_2.BP_{err}$, where γ is the learning rate. In general, the value of learning rate is between 0.2 and 0.5.

Step 8: The hidden layer outputs are computed as:

$$O_j^h = \frac{1}{1+e^{-a\sum_1^m W_j^i x_i}} \tag{4}$$

Where χ_i, the net input to the ith input unit; O_j^h, the output of the jth hidden layer neuron and W_j^i is the weight on the connection from the ith input unit to the jth hidden unit. The actual output kth hidden layer is:

$$O_k^h = \frac{1}{1+e^{-a\sum_1^n W_k^j U_j}} \tag{5}$$

He actual output for the kth unit and W_k^j is the connection weight from the jth hidden unit to kth output unit. The error term between output layer neuron and hidden layer neuron is:

$$\delta_k^o = O_k(1 - O_k)(T_k - O_k) \tag{6}$$

Where, T_k is the target output for the kth output unit and δ_k^o is the error term for the kth output unit. δ_j^h is the signal error for the jth hidden unit, given as:

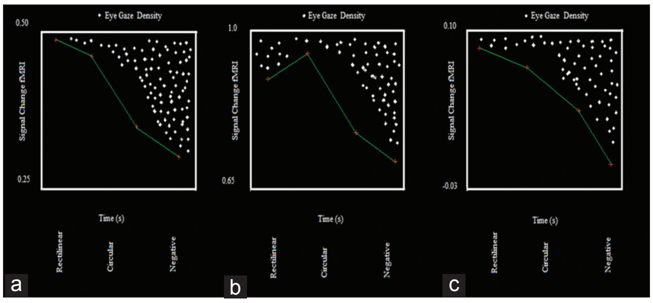

Fig. 2: (a-c) Stimuli results for arrays of rectilinear shapes with rounded and negative images while depicting the eye gaze density for each of the images. Thus, from this viewpoint the procedurally created viewership of the certain images with specific properties of being rectilinear or curvilinear or negative object images that there is significant bias for the rectilinear images than with the circular or rounded or curvilinear images. Finally, the negative images did not preserve the object size thus having hard time recognizing those images by the subject

Table 1: Power of stimuli images in order of difficulty in detection by segmenting objects in the image as found in the experiment

Experiment	Order of difficulty in detection
1	Circle>Cones>Triangle>Square
2	Round>Rectangle
3	Spheres>Cones>Pyramids>Cubes
4	Circles>Dodecagons>Hexagon>Triangles>Square
5	Circles>Triangles>Square

$$\delta_j^h = O_j(1 - O_j)\sum_{k=1}^{p} \delta_k^o W_{jk} \tag{7}$$

The weights on the output layer are adjusted by:

$$W_{jk}^o(t+1) = W_{jk}^o(t) + (k.\delta_k^o.O_j) \tag{8}$$

Where, η is learning rate of output layer. The weights of hidden layer neurons are adjusted as:

$$W_{jk}^h(t+1) = W_{ij}^o(t) + (j.\delta_k^o.X_i) \tag{9}$$

Final error is calculated (if $(T_k - O_k) \geq 0$):

$$\delta_k^o = (1 + e^{(T_k - O_k)^2})O_k.(1 - O_k) \tag{10}$$

Otherwise

$$\delta_k^o = -(1+e^{(T_k - O_k)^2})O_k.(1-O_k) \quad (11)$$

Now, the rules are generated in the form of fuzzy error classifiers and these fuzzy rules are generated to recreate the recognition process and train the artificial neural system for visual recognition. Major steps of the system are:

- Repeat the process until the back propagation error is minimized as BP_{err} <0.1.
- Check for next segmented objects and repeat until recognition of all object is over.
- The average error is computed for all objects which are in correlation with other segmented ones a supposed to be less than 40% the total error.
- Finally, the above process has to be repeated till specified number of epochs.
- Once error threshold is reached, the object recognized is displayed.

RESULTS AND DISCUSSION

Here, as shown in Fig. 1 in the Ventral view of the averaged brain for the localization of relational activity evoked in response with rectilinearity versus circular real-world objects and faces, of the subjects in the experiment. The yellowish region of the brain areas represents the magnified antirational responses resultant from rectilinear versus circular visual object stimuli.

The imaging results revealed marked neural response which tends to serve to complement the behavioral findings of the significant restoration of the ability to recognize for such visual images. Altogether, these causes to give an explanation on the long-standing question of why photographic negatives are hard to recognize. These results suggest that the difficulty in analyzing negative images is driven in significant parts by dissolving the geometric shapes in 2D contrast polarity relations between the essential regions of the face defined by a combination of rectilinear and circular shaped objects. The special

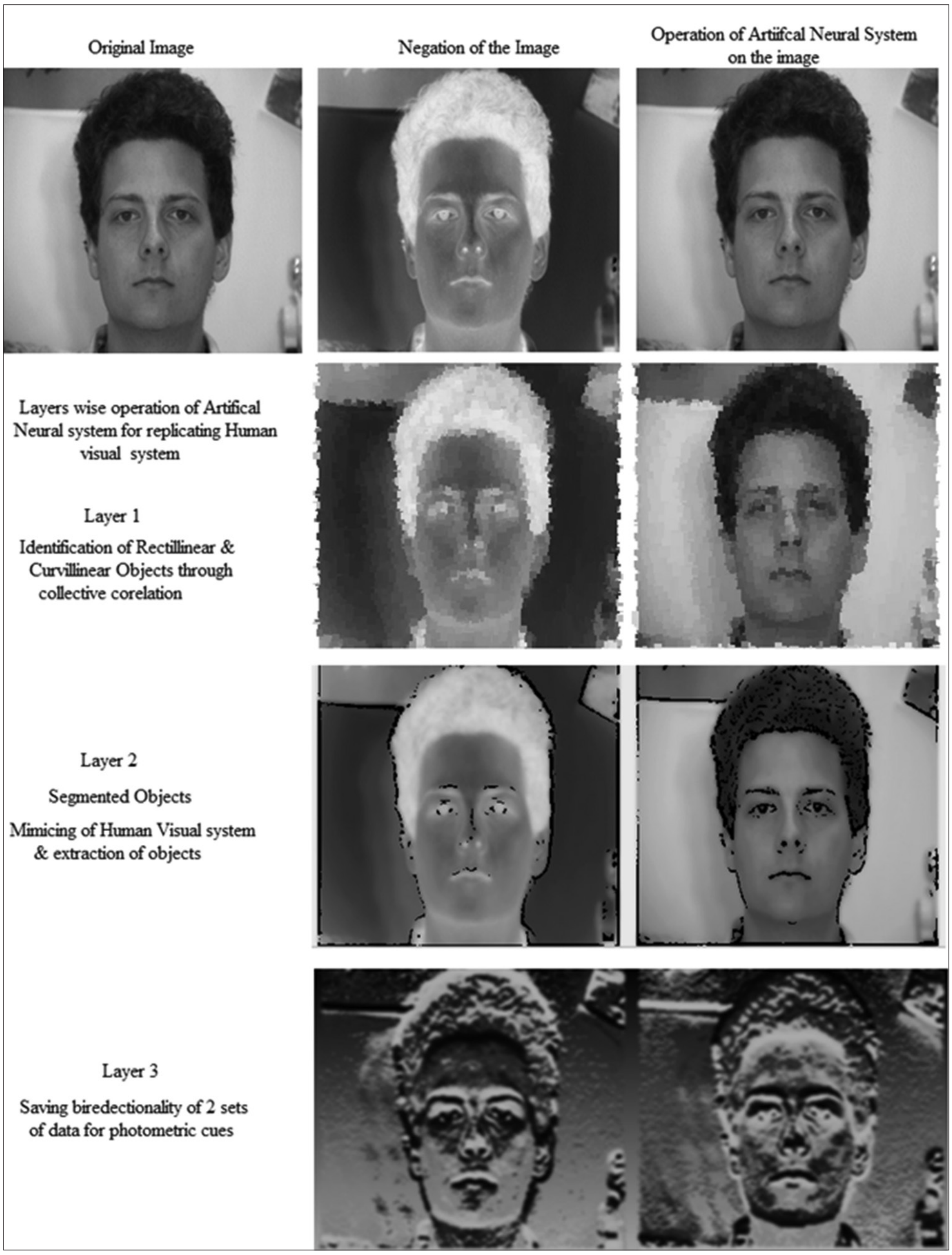

Fig. 3: In process instances of the artificial neural system replicating the human visual system computationally based on the visual stimuli and segmenting objects of a grey-image and negation of the original image; for feature extraction and object recognition

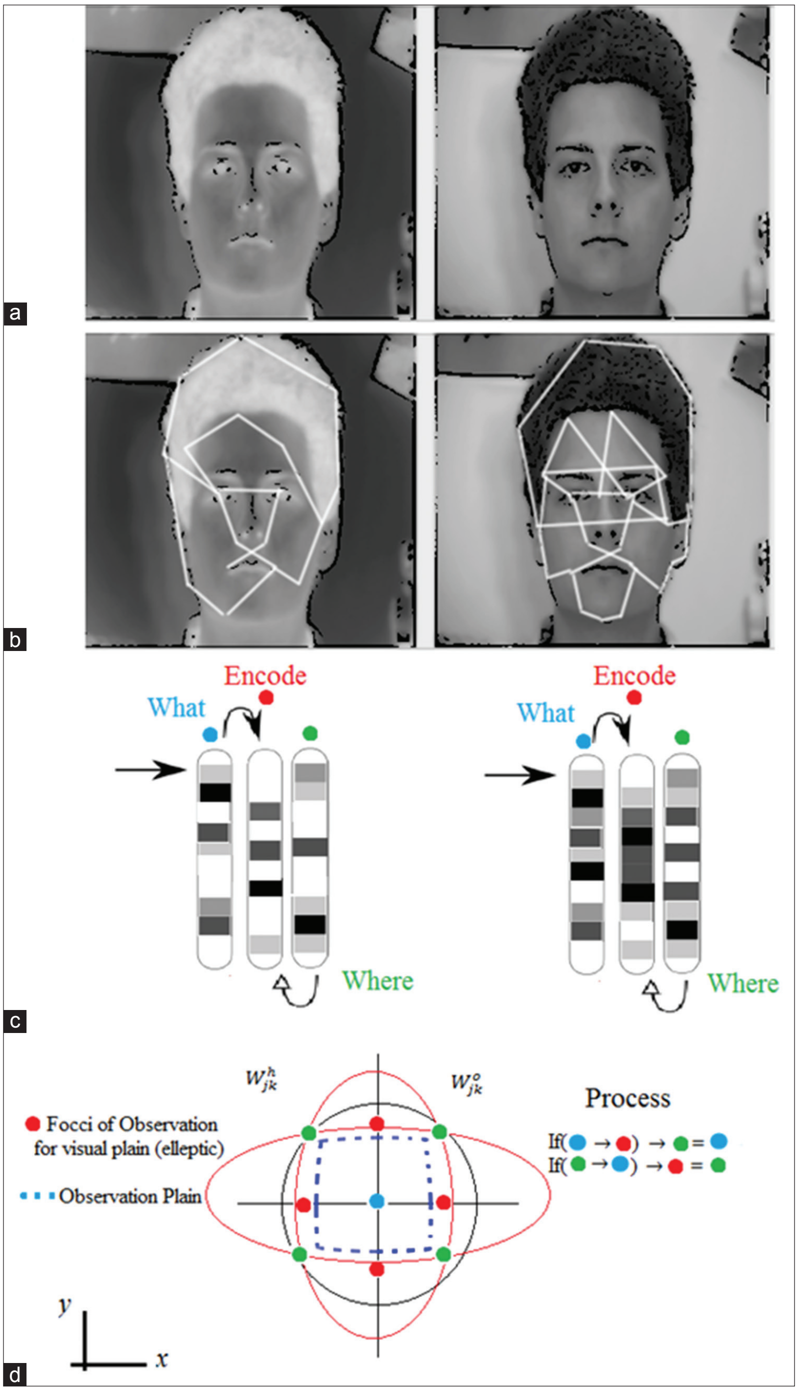

Fig. 4: Mechanism of Human Visual System detected by Artificial Neural System. (a) Segmented Image by the Subjects, (b) track of eye-gaze movements of the subjects for the negative and gray scale image, (c) memory stacks generated from data recorded by artificial neural system for weighted imagery regions by the subjects while recording the features of the image during face recognition for both gray scale version and its negative. As shown in the memory stacks are divides in to 3 units for the specific image which represents what the observer is looking, where the observer is looking and the weighted encoded features in relational to the what and where the image sections are most prominent. Thus, from the memory stacks the more white spaces in the negative of the given image inferred that why it's difficult for the subjects to recognize it than with the actual image, which shows more prominent scales of colored spaces in relational with the what and where before features are being encoded, (d) the plain shows the computational logic behind tracking of the eye movements by the subjects based on data derived from the artificial neural system. Here, the two elliptical plain shows the observational plains from each eyes and thus the intersection of the two plains gives the square-curved plane of main observational plain for which other than outside of it is blurred for the subjects. That results in more prone of the subject to identify rectilinear shapes than the curvilinear ones. Thus, the logic derived from data of artificial neural system drives the observational plain for encoding features are: (i) Rule for encoding features: If what observer is looking at is judge for encoding for its property; when, where the observer will be looking for encoding of features is changed to what it shall looked for observation from the four corners of the observational plain (represented by green color dots) which should be selected based on intersection of those green colored positional observational dots with the lines or curves of the image, (ii) rule for movement of gaze: If where the observation should look next for what is to observe is to be encoded when the next features to be encoded is changed to where the observation dots point out next location based on weighted colored red encoded dots intersects with the curves or lines coinciding with it

significance of eye gazes is that it perceives objects differently in the collection of segmented order (for a fractal image) overlaid in the neighborhood which gives out the data from the same experiment conducted with different diameters of the circular aperture (Fig. 2). Thus, this finding explains the perceptual significance of reconciliation of photometric relationships with human's ability to identify and recognize line-drawings of faces rather easily than circular ones. Evidently, line drawings mainly contain contour information and very little photometric information which is held responsible for defining luminance relation. However, when it comes to facial recognition or of fractal images the density and weight of the lines affects the relative intensity of different regions. Thus, the contour lines included in such depictions corresponds not only to low-level edge maps instead it embodies the images' photometric structure. It is the skillful inclusion of these photometric cues by the overlapping of contour lines which in our experiment make the human subjects more prone to easily recognize line-drawings which are latter replicated by the computer algorithm as shown in Fig. 3.

Layer 1 shows the selective identification of objects, that is, mapping of digital image data by the artificial system based on the data derived from subjects under experimentation by sensing the crowd localization through the collective correlation between the pixels position, photometric cues, and time slot for eye gaze invested on it for both negative image and gray-scale image, whereas in the Layer 2, the curvilinear objects are segmented by artificial neural system based on same rule of observation for the subjects, since the artificial neural system depends on localization of eye gaze. This shows us the vast amount of different degrees of segmentation of essential features and proves us why it is difficult for the humans to identify faces for negation of the image; thus, it represents the most prominent features identified for face recognition by the subjects as the similar pattern is mimic by the artificial neural system. While in Layer 3, the image represents the perseverance of bidirectionality of two sets of photometric cues (on further iteration past the mimicking period) within the same imagery by the artificial neural system, which is impossible for the human visual system (which in future study will help frame a better visual system for robotics and artificial beings). Thus, inferring from the above results the human visual system for visual encoding depends on two factors namely:

i. What we are looking at;
ii. Where we are looking at.

This effect is simultaneously exhibited with the subjects and explained by the artificial neural system in Fig. 4.

The extraction of ordinal by the human visual system is achieved by contour response profiles of neurons in the primary visual cortex; as neurons of V1 exhibits rapid saturating responses which are a function of contour lines. These profiles are approximations of step-functions as in our modeled artificial neural system, which characterize the role of an ordinal comparator. Thus, neurons of the mammalian visual pathway give a plausible substrate for extracting collective ordinal image relationships; even that includes overlaid images. Negation effects are significantly more pronounced for faces than for other classes of objects because face constitutes a homogenous class with a consistent set of photometric relationships within them. Given this invariability, an ordinal code would privilege one direction of illumination over the other. Such an encoding will be affected by negation.

CONCLUSION

The results pertain the relevance of our study to specific criteria of visibility in autism to filter out the results. In neurological disorders such as autism is associated with face processing abnormalities as the individuals with autism are known to avoid eye-contact and tend to focus more on the regions of the mouth. Thus, the identification relationships are weighted more toward the mouth, then we expect that mouth chimeras is more facilitate toward enhanced recognition performance of autistic observers rather than that of eye-chimeras, in contrast to the patterns of results we have described above with neuro-typical observers. Thus, the nature of facial representation in autism is summarized as per our findings that the potential answer to the long-standing question of why faces are hard to recognize in negative contrast images. The rectilinear sensitivity is developed more quickly, within each person's lifetime which forms the selectivity at the level of fMRI. In general, they suggest that contrast polarity relations between face regions in the vicinity of the eyes and mouth is embodied in the visual system's facial representations and serve as strong determinants of recognition performance. In PPA, even very abstract shape differences mainly cubes versus spheres forms differences in fMRI amplitude because is comparable with that produced by a well-known category based such as faces vs scenes contrast.

REFERENCES

1. Galper RE. Recognition of faces in photographic negative. Psychon Sci 1970;19:207-8.
2. George N, Dolan RJ, Fink GR, Baylis GC, Russell C, Driver J. Contrast polarity and face recognition in the human fusiform gyrus. Nat Neurosci 1999;2(6):574-80.
3. Hayes T, Morrone MC, Burr DC. Recognition of positive and negative band pass filtered images. Perception 1986;15(5):595-602.
4. Nederhouser M, Yue X, Mangini MC, Biederman I. The deleterious effect of contrast reversal on recognition is unique to faces, not objects. Vision Res 2007;47(16):2134-42.
5. Phillips RJ. Whyare faces hard to recognize in photographic negative. Percept Psychophys 1972;12(5):425-6.
6. White M. Effect of photographic negation on matching the expressions and identities of faces. Perception 2001;30(8):969-81.
7. Bruce V, Young A. In the Eye of the Beholder: The Science of Face Perception. Oxford: Oxford University Press; 1998.
8. Cavanagh P, Leclerc YG. Shape from shadows. J Exp Psychol Hum Percept Perform 1989;15(1):3-27.
9. Hill H, Bruce V. Effects of lighting on the perception of facial surfaces. J Exp Psychol Hum Percept Perform 1996;22(4):986-1004.
10. Johnston A, Hill H, Carman N. Recognising faces: Effects of lighting direction, inversion, and brightness reversal. Perception 1992;21(3):365-75.
11. Kemp R, Pike G, White P, Musselman A. Perception and recognition of normal and negative faces: The role of shape from shading and pigmentation cues. Perception 1996;25(1):37-52.
12. Moscovitch M, Winocur G, Behrmann M. What is special about face recognition? Nineteen experiments on a person with visual object agnosia and dyslexia but normal face recognition. J Cogn Neurosci 1997;9(5):555-604.
13. Beymer D, Poggio T. Image representations for visual learning. Science 1996;272(5270):1905-9.
14. Turk M, Pentland A. Eigenfaces for recognition. J Cogn Neurosci 1991;3(1):71-86.
15. Bruce V, Langton S. The use of pigmentation and shading information in recognising the sex and identities of faces. Perception 1994;23(7):803-22.
16. Vuong QC, Peissig JJ, Harrison MC, Tarr MJ. The role of surface pigmentation for recognition revealed by contrast reversal in faces and greebles. Vision Res 2005;45(10):1213-23.
17. Braje WL. Illumination encoding in face recognition: Effect of position shift. J Vis 2003;3(2):161-70.
18. O'Toole AJ, Jonathon Phillips P, Jiang F, Ayyad J, Penard N, Abdi H. Face recognition algorithms surpass humans matching faces over changes in illumination. IEEE Trans Pattern Anal Mach Intell 2007;29(9):1642-6.
19. Rai A. Attribute based level adaptive thresholding algorithm for object extraction. J Adv Robotics 2015;1(2):64-8.
20. Rai A. Dynamic data flow based spatial sorting method for GPUs: Software based autonomous parallelization. Recent Trends Parallel Comput 2014;1(1):15-8.
21. Rai A. Characterizing face encoding mechanism by selective object pattern in brains using synthetic intelligence and its simultaneous replication of visual system that encode faces. Res Rev J Comput Biol 2014;3(2):1-8.