International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

An Efficient Method for Object Recognition Using Histogram

Dr. Sandeep. C. S[1], Aswathy J[2], Ajeena A[3]

[1]Associate Professor, Dept. of ECE, Jawaharlal College of Engineering and Technology, E.mail: dr.sandeep.cs@gmail.com
[2]Assistant Professor, Dept. of ECE, Jawaharlal College of Engineering and Technology aswaty1989@gmail.com
[3]Assistant Professor, Dept. of ECE, Jawaharlal College of Engineering and Technology, ajeena4093.ece@jawaharlalcolleges.com

ABSTRACT –

A simple and effective object recognition scheme is to represent and match images on the basis of color histograms. To obtain robustness against varying imaging circumstances (e.g. a change in illumination, object pose, and viewpoint), color histograms are constructed from color invariants. However; in general, color invariants are negatively effected by sensor noise due to the instabilities of these color invariant transforms at many RGB values. To suppress the effect of noise blow-up for unstable color invariant values, in this paper; color invariant histograms are computed using variable kernel density estimators. To apply variable kernel density estimation in a principled way, models are proposed for the propagation of sensor noise through color invariants. As a result, the associated uncertainty is known for each color invariant value. The associated uncertainty is used to derive the parameterization of the variable kernel density estimator during histogram construction.

We aim at combining color and shape invariants for indexing and retrieving images. To this end, color models are proposed independent of the object geometry, object pose, and illumination. From these color models, color invariant edges are derived from which shape invariant features are computed. Computational methods are described to combine the color and shape invariants into a unified high-dimensional invariant feature set for discriminatory object retrieval. Experiments have been conducted on a database consisting of 500 images taken from multicolored man-made objects in real world scenes. From the theoretical and experimental results it is concluded that object retrieval based on composite color and shape invariant features provides excellent retrieval accuracy. Object retrieval based on color invariants provides very high retrieval accuracy whereas object retrieval based entirely on shape invariants yields poor discriminative power. Furthermore, the image retrieval scheme is highly robust to partial occlusion, object clutter and a change in the object's pose.

***Index Terms*— Object recognition, noise robustness, histogram, construction, noise propagation, kernel density estimation, matching., color invariants, shape invariants.**

I. Introduction

Color provides powerful information for object recognition. To provide object recognition robust against the accidental imaging conditions (e.g. illumination, shading, highlights, and viewpoint), color histograms are computed from color invariants. For example, simple and effective illumination-independent color ratio's have been proposed by Funt and Finlayson .Further, for the dichromatic reflection model, Gevers and Smeulders proved that normalized color *rgb(c1c2c3)* to a large extent invariant to a change in camera viewpoint, object pose, and for the direction and intensity of the incident light [1]. In addition, the hue color is insensitive to highlights under white light or white-balanced camera. However, these color invariant transforms bring them various serious drawbacks since these transformations are singular at some points and unstable at many others. For example, color ratio's, *rgb* and c1c2c3 color space become unstable near the black point while hue *H* and $1_1 1_2 1_3$ are very unstable along the achromatic axis .As a consequence, a small perturbation of *RGB* values will cause a large jump in the transformed values introducing severe errors during histogram construction. Traditionally, the effects of noise blow-up at unstable colors are suppressed by ad hoc thresholding. For example, all pixels in a color image are discarded with a local saturation and intensity smaller then 5 percent of the total. Inevitably, more elaborated computational methods are required to construct robust histograms from color invariants.

Most of the work on shape-based object recognition rely on matching sets of local image features (e.g., edges, lines and corners) to three-dimensional (3-D) object models invariant to geometric transformations (e.g., translation, rotation, scale, and affine transformation) and significant progress has been achieved. As an expression of the difficulty of the general problem, most of the geometry- based matching schemes can handle only simple, flat, and rigid man-made objects. Shape features are rarely adequate for discriminatory object recognition of 3-D objects from arbitrary viewpoints in complex scenes [2]. It shows that hue and normalized color are singular at some *RGB* values and unstable at many others. For instance, the essential singularity of normalized coordinates is at black (R = G = *B* = O), whereas for the hue the singularity is at the achromatic axis *(R = G = B).* As a consequence, both color spaces become unstable near the singularity where a small perturbation of *RGB* value might cause a large jump in the transformed values. Traditionally, these effects are either ignored or suppressed by ad hoc thresholding of the transformed values. For example, it rejects normalized color values if the sum of *RGB* is less than 30, and rejects hue values if the saturation times the intensity is less than 9. Healey rejects all sensor measurements that fall within the sphere of radius 4σ centered at the origin in *RGB* space. In order to indicate how the color space transform influences the mean, variance and covariance of the colors under the influence of noise, Burns and Berns analyze the error propagation from a measured color signal to the CIE *L*a* b** color space. Shafarenko, Petrou and Kittler use an adaptive filter for noise reduction in the CIE *L*u*v** space prior to 3-D color histogram construction. The filter width depends on the covariance matrix of the noise distribution in the CIE *L*u*v** space. As opposed to previous work, the aim of this paper is to use kernel density estimators to construct color invariant histograms. To apply variable kernel density estimation in a principled way, the error propagation of sensor noise is analyzed through the color invariants. As a result, the associated uncertainty is known for each color invariant value. Then, kernel sizes are adapted with respect to the amount of noise blow-up for unstable color invariant values. As a consequence, unstable color invariant values will contribute less to the final histogram than stable color invariant values. Although our method of variable kernel density estimation is suited for different color invariants,we focus on normalized color *rgb* **a**nd hue as these color spaces are widely in use in computer vision tasks [3].

As opposed to shape information, other retrieval schemes are entirely on the basis of color. Swain and allard made a significant contribution in introducing color for object search. Based on the opponent color model, they show that image retrieval based on histogram matching is to a large degree robust to changes in object pose and shape. The histogram based matching scheme is extended by Funt and Finlayson and Nayar and Bolle to make the method illumination independent by indexing on color ratio's computed from neighboring image points. However, the color ratio's are negatively affected by the geometry of the object [4].

In this paper, we want to arrive at combining color and shape invariants for the purpose of image indexing and retrieval. To that end, a retrieval scheme is proposed making use of local color invariant information to produce semi global shape invariants to obtain a viewpoint invariant, high-dimensional object descriptor to be used as an index for discriminatory image retrieval. To achieve this, color invariant features are proposed according to the following criteria: invariance to the viewpoint, geometry of the object, and illumination conditions. Then, from these color models, color invariant edges are derived from which the shape features are computed. Shape features are independent up to a change in viewpoint (i.e., projective transformation). Computational methods are proposed to combine color and shape invariants into a unified high-dimensional invariant feature space. The image retrieval scheme is designed according to the following criteria: high discriminative power, and robustness against fragmented, occluded and overlapping objects.

2. BLOCK DIAGRAM REPRESENTATION

Fig1 shows the Object Recognition Scheme. A query image is taken from the database images and it is recorded at different illumination conditions. Next step is to extract the color and shape invariants. Then color invariants are estimated using kernel density estimators. After that the angle calculation is made for shape invariants. Then histograms are constructed depends on the calculations made for color and shape invariants. Next step is to find the Euclidean distance measurement between the sampled image and the image in the data base. Based on the distance measured, the images are classified.

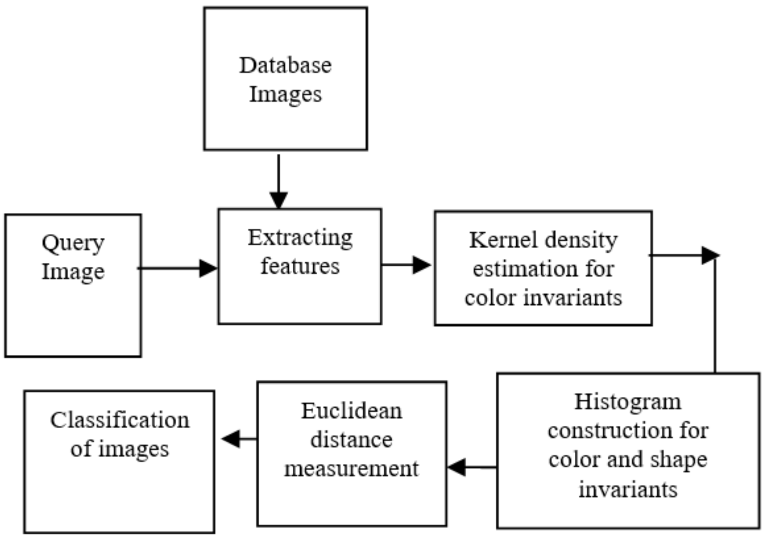

Fig 1 Object Recognition Scheme

3. FEATURES EXTRACTION

This means to extract the color and shape invariants such as normalized color space, opponent color space, and hue of the image, edges, lines and corners

3.1 Color Invariants

A color perceived by the human eye can be defined by a linear combination of the three primary colors red, green and blue. These three colors form the basis for the RGB-color space. Hence, each perceivable color can be defined by a vector in the three-dimensional color space. The intensity is given by the length of the vector, and the actual color by the two angles describing the orientation of the vector in the color space.

Some simple way to achieve chromaticity invariance with respect to illumination changes is to normalize the tristimulus coordinates, such that they sum to 1. Based on the measured RGB-values, the normalized color *rg* is computed by:

$$r = R/(R+G+B)$$

$$g = G/(R+G+B)$$

where *rg* is a color invariant for matte surfaces by factoring out dependencies on illumination and object geometry and, hence, only dependent on the sensors and the surface albedo.

Further, we focus on the opponent color space defined as a linear transformation of RGB that matches the physiology of the human visual system. The three coordinates are a achromatic value and two chromatic coordinates given by green-red and blue-yellow differences.

$$o_1(R,G,B) = (R-G)/2$$

$$o_2(R,G,B) = (2B-R-G)/4$$

The opponent color space is well-known and has its fundamentals in human perception. The opponent color space o1o2 is independent of highlights .Equal argument holds for o_2. Note that object geometry and o1o2 is still dependent on and consequently, being sensitive to shading. [5]

Next we move on to hue which is defined as the property of colors by which they can be perceived as ranging from red through yellow, green, and blue, as determined by the dominant wavelength of the light. A color system, or model, measures color by hue, saturation and luminance. The hue is the predominant color, the saturation is the color intensity, and the luminance is the brightness.

The hue θ is computed as

$$\theta = \arctan\left(\frac{\sqrt{3}(G-B)}{(R-G)+(R-B)}\right)$$

It is also insensitive to illumination, object geometry, and highlights Additive Gaussian noise is widely used to model thermal noise and is the limiting behavior of photon counting noise and film grain noise. Therefore, in this paper, we assume that sensor noise is normally distributed for random errors, it is used as an universal formula for various kinds of errors [6].

Uncertainty for the normalized coordinates is computed as

$$\sqrt{\quad} \qquad (R+G+B)$$

Assuming normally distributed random quantities ,the standard way to calculate the standard deviations σR, σG, and σB is to compute the mean and variance estimates derived from homogeneously colored surface patches in an image under controlled imaging conditions. From the analytical study, it can be derived that normalized color becomes unstable around the black point R =G =B =0 [7].

The uncertainties of o1 and o2 are given by

$$\sigma_{o_1} = \frac{1}{2}\sqrt{\sigma_G^2 + \sigma_R^2}$$

$$\sigma_{o_2} = \frac{1}{4}\sqrt{4\sigma_B^2 + \sigma_G^2 + \sigma_R^2},$$

which are the same (stable) at all RGB points.

The uncertainty for the hue is given by

$$\sigma_\theta = \sqrt{\frac{3}{4}\frac{(B-G)^2\sigma_R^2 + (B-R)^2\sigma_G^2 + (R-G)^2\sigma_B^2}{(R^2+G^2+B^2-RG-RB-GB)^2}},$$

which is unstable at low intensity and saturation (i.e., the gray axis R =G =B).

In conclusion, it can be analytically derived that normalized color is unstable at low intensity. Hue is unstable at low intensity and saturation. Opponent color is relatively stable at all RGB values.

3.2 Shape Invariants

In this section, shape invariants are discussed measuring geometric properties of a set of coordinates of an image object independent of a coordinate transformation. We discuss similarity and projective invariants.

A. Similarity Invariant

For image locations (x_1,y_1), (x_2, y_2) and (x_3, y_3), $g_E()$ is defined as a function which is unchanged as the points undergo any two-dimensional (2–D) translation, rotation and scaling transformation, yielding the well-known similarity invariant:

$g_E((x_1,y_1), (x_2, y_2), (x_3, y_3)) = \theta$

where θ is the angle at image coordinate (x_1,y_1) between line $(x_1,y_1)(x_2, y_2)$ and $(x_1,y_1)(x_3, y_3)$.

B. Projective Invariant

For the projective case, geometric properties of the shape of an object should be invariant under a change in the point of view. From the classical projective geometry we know that the so called cross-ratio is independent of the projection viewpoint [8]. we derive the projective invariant $g_P(\,)$ defined as

$$g_P((x_1, y_1), (x_2, y_2), (x_3, y_3), (x_4, y_4), (x_5, y_5)) = \frac{\sin(\theta_1+\theta_2)\sin(\theta_2+\theta_3)}{\sin(\theta_2)\sin(\theta_1+\theta_2+\theta_3)}$$

where θ_1, θ_2, θ_3 are the angles at image coordinate (x_1,y_1), between $(x_1,y_1)(x_2, y_2)$ and $(x_1,y_1)(x_3, y_3)$, $(x_1,y_1)(x_3, y_3)$ and $(x_1,y_1)(x_4, y_4)$, (x_1,y_1), (x_4, y_4) and $(x_1,y_1)(x_5, y_5)$ respectively.

4. Histogram CONSTRUCTION

Histogram is a charting format that displays horizontal or vertical bars. The lengths of the bars are in proportion to the values of the data items they represent.

Many digital cameras use histograms to display the brightness of the image. Using 256 vertical bars to represent brightness levels from 0 to 255, the bar on the leftmost side of the chart is the darkest pixel level (0), and the rightmost bar is the lightest (255). The height of the bars represents the relative number of pixels in the image that contain that brightness level [9, 10].

4.1 Color Invariant Histogram Formation

A density function ***f*** gives a description of the distribution of the measured data. Perhaps the best known density estimator is the histogram. The (one-dimensional) histogram is defined as

$f(x) = \frac{1}{nh}$ (number of X_i , in the same bin as *x)*

where *n* is the number of pixels X_i, in the image, *h* is the bin width and ***x*** the range of the data. Two choices have to be made when constructing a histogram. First, the binwidth parameter needs to be chosen. Secondly, the position of the bin edges needs to be established. Both choices effect the resulting estimation [11].

Alternatively, the kernel density estimator is insensitive to the placement of the bin edges.

$$\hat{f}(x) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right).$$

In the *variable* kernel density estimator, the single *h* is replaced by *n* values.*(Xi)**i*** ,= 1, . . . *n* [12]. This estimator is of the form

$$\hat{f}(x) = \frac{1}{n}\sum_{i=1}^{n} \frac{1}{\alpha(X_i)} K\left(\frac{x - X_i}{\alpha(X_i)}\right).$$

The kernel centered on *Xi* , has associated with it its own scale parameter *a* (X_i) , thus allowing different degrees of smoothing. To use variable kernel density estimators for color images, we let the scale parameter be a function of the RGB-values and the color space transform [13].

Assuming normally distributed noise, the distribution is approximated well by the Gauss distribution

$$K(x) = \frac{1}{\sqrt{2\pi}} \exp^{\frac{-x^2}{2}} .$$

Then, the variable kernel method estimating the univariate, directional hue density is as follows:

$$\hat{f}(\theta) = \frac{1}{n}\sum_{i=1}^{n} \sigma_{\theta_i}{}^{1} K\left(\frac{(\theta - \theta_i) \bmod (\pi)}{\sigma_{\theta_i}}\right)$$

The variable kernel method for the bivariate normalized rg kernel is given by

$$\hat{f}(r,g) = \frac{1}{n}\sum_{i=1}^{n} \sigma_{r_i}^{-1} K\left(\frac{r - r_i}{\sigma_{r_i}}\right) \sigma_{g_i}^{-1} K\left(\frac{g - g_i}{\sigma_{g_i}}\right)$$

Similarly, the variable kernel method for the bivariate normalized o_1o_2 kernel is given by:

$$\hat{f}(o_1, o_2) = \frac{1}{n}\sum_{i=1}^{n} \sigma_{o_{1_i}}^{-1} K\left(\frac{o_1 - o_{1_i}}{\sigma_{o_{1_i}}}\right) \sigma_{o_{2_i}}^{-1} K\left(\frac{o_2 - o_{2_i}}{\sigma_{o_{2_i}}}\right)$$

4.2 Shape Invariant Histogram Formation

In this section, shape invariant histograms are constructed. We use $l_4l_5l_6$ - based color invariant edges as feature points. These edges are viewpoint-independent, discounting shading, illumination intensity and direction, shadows and highlights [14]. A one-dimensional (1–D) histogram is constructed in a standard way on the angle axis expressing the distribution of angles between color invariant edge triplets mathematically specified by

$$\mathcal{H}_D(i) \hat{=} \eta(g_E(\vec{x}_1, \vec{x}_2, \vec{x}_3) = i)$$

only computed for $\overrightarrow{x_1}$ ¹ $\overrightarrow{x_2}$ ¹ $\overrightarrow{x_3}$ Î E^{Ik}, where E^{Ik} is the set of edge maxima Thus, between each triplet of color edge maxima, the angle denoted by is computed and used as an index [15]. Hence, each particular bin sum can be seen as the number of color edge triplets generating the same angle. In a similar way, a 1-D histogram is defined on the cross ratio axis expressing the distribution of cross ratios between color edge quintets

$$\mathcal{H}_E(i) \hat{=} \eta(g_P(\vec{x}_1, \vec{x}_2, \vec{x}_3, \vec{x}_4, \vec{x}_5) = i)$$

5. Distance Measure

The Euclidean distance or Euclidean metric is the "ordinary" distance between the two points. The distance between the image in the database and sample image are calculated for the classification of images using this approach.

The Euclidean distance between the two n dimensional (row or column) vectors x and y is defined as the scalar [16,17].

d (x, y)=||x - y||= ||y - x||

=[(x1-y1)2 +………….+ (xn-yn)2]½

6. RESULTS

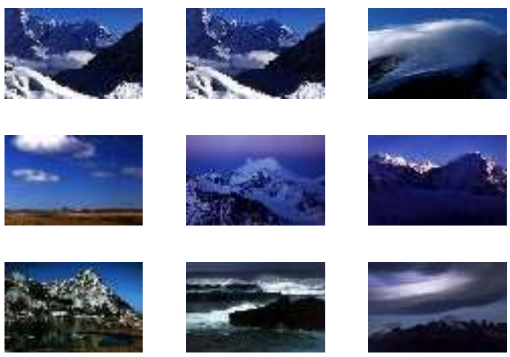

Fig.2 Query Image and retrieved Images

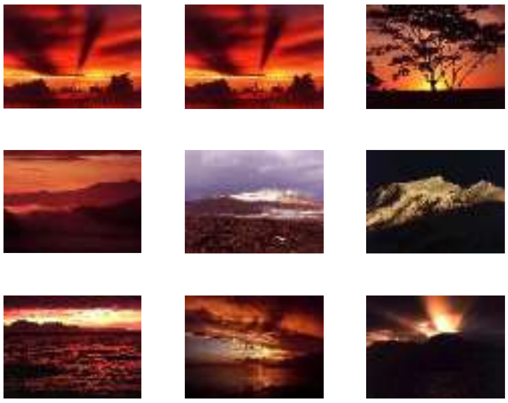

Fig.3 Query Image and retrieved Images

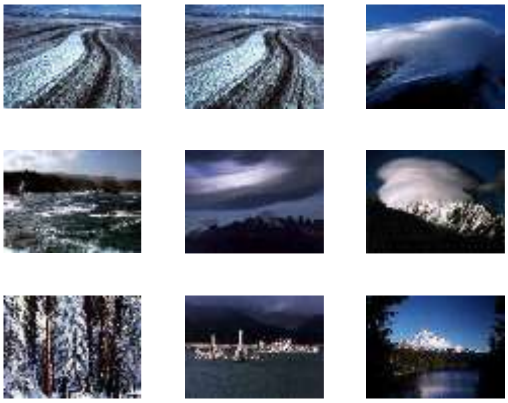

Fig.4 Query Image and retrieved Images

CONCLUSION

In this paper, variable kernel density estimation is used to construct robust color and shape invariant histograms. The shape features are extracted from the classical geometry. The variable kernel density estimation is derived from a theoretical framework for noise propagation through color invariants. In this way, the associated uncertainty is computed for each color invariant value, which is used to steer the kernel sizes. From the theoretical and experimental result, we conclude that kernel density estimator overcome the problem of adhoc thresholding at unstable color invariants. Further, our method is less sensitive to Gaussian noise than traditional histogram construction schemes. A drawback of the variable kernel method compared to traditional histogram construction is that the method is computationally more expensive.

References

[1]. Li, P., Chen, X. and Shen, S. (2019) Stereo R-CNN Based 3D Object Detection for Autonomous Driving. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, 15-20 June 2019, 7636-7644.

[2]. Zhang, X., Yi, W.-J. and Saniie, J. (2019) Home Surveillance System Using Computer Vision and Convolutional Neural Network. 2019 IEEE International Conference on Electro Information Technology (EIT), Brookings, 20-22 May 2019, 266-270.

[3]. Sandeep C.S., Vijayakumar N., Sukesh Kumar A., “A Novel Scheme for the Early Diagnosis of Alzheimer’s Disease through MRI Analysis”, ADBU Journal of Engineering Technology, ISSN: 2348-7305, vol. 9, no. 1, pp. 1-5, 2020.

[4]. Zhang, R., Shao, Z., Huang, X., Wang, J. and Li, D. (2020) Object Detection in UAV Images via Global Density Fused Convolutional Network. Remote Sensing, 12, Article No. 3140.

[5]. Sandeep C. S., Sukesh Kumar A., K. Mahadevan, Manoj P., “Analysis of Retinal OCT Images for the Early Diagnosis of Alzheimer's Disease”, Springer-Advances in Intelligent Systems and Computing book series, ISBN: 978-3-319-74807-8, vol.749, pp. 509-520, 2019.

[6]. Girshick, R., Donahue, J., Darrell, T. and Malik, J. (2013) Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, 23-28 June 2014, 580-587.

[7]. Brian V. Funt and Graham D.Finlayson,"Color Constant Color Indexing", *IEEE Transactions on Pattern Analysis and Machine Intelligence*,, vol. 17, no. 5 may 1995

[8]. Sandeep C.S., Vijayakumar N., Sukesh Kumar A., "A Novel Approach for the Early Diagnosis of Alzheimer's Disease", ACTA TECHNICA CORVINIENSIS – Bulletin of Engineering, ISSN: 2067–3809, TOME XIII, FASCICULE 4, 2020

[9]. Jan-Mark Geusebroek Dennis Koelma Arnold W.M. Smeulders Theo Gevers,"Image Retrieval and Segmentation based on Color Invariants", *IEEE* 2000

[10]. Arnold W M. Smeulders and Jan-Mark Geusebroek and Theo Gevers,"Invariant Representation in Image Processing *".IEEE* 2001

[11]. Joost van de Weijer, ,Theo Gevers, andJan-Mark Geusebroek," Edge and Corner Detection by Photometric Quasi-Invariants" *IEEE Transactions on Pattern Analysis and Machine Intelligence,,*vol. *27, no.* 4, april 2005

[12]. J. van de Weijer Th. Gevers ," Boosting Saliency in Color Image Features ",IEEE 2005

[13]. Sandeep C. S., Sukesh Kumar A., K. Mahadevan, Manoj P., "Breakthrough in the Early Diagnosis of Alzheimer's Disease", International Journal of Computer Science and Programming Language, vol. 6, no. 1, pp. 1-12, 2020.

[14]. Joost van de Weijer, Theo Gevers, andAndrew D.Bagdanov," Boosting Color Saliency in Image Feature Detection", *IEEE Transactions on Pattern Analysis and Machine Intelligence*,vol. 28, no. 1 , january 2006

[15]. Sandeep C. S., Sukesh Kumar A., K. Mahadevan, Manoj P., "Early Prediction of Alzheimer's Disease Using OCT Imaging Technique", Journal of Alzheimer's Research and Therapy, vol.1, no.1,

[16]. Xiaolin Wu, Senior Member, IEEE, and Nasir Memon,,"Context-Based, Adaptive, Lossless Image Coding",*IEEE Transactions on Communications* vol. 45,no 4,april 1997

[17]. Sandeep C.S., Vijayakumar N., Sukesh Kumar A., "Can we predict Alzheimer's Disease through the eye lens?", International Journal of Clinical and Experimental Ophthalmology, ISSN: 2577-140X, vol. 4, no. 1, pp. 38-40, 2020.