PERFORMANCE EVALUATION OF OBJECT RECOGNITION USING SKELETAL SHOCK GRAPH: CHALLENGES AND FUTURE PROSPECTS

Pankaj Manohar Pandit\textsuperscript{1}, Sudhir Gangadhara Akojwar\textsuperscript{2}
\textsuperscript{1}Dept. of Ele. & Tele. Engg., Jawaharlal Darda Institute of Engg. & Technology, Yavatmal  
\textsuperscript{2}Principal, Dr. Bhausahaeb Nandurkar College of Engg. & Technology, Yavatmal

\section*{Abstract}
This paper describes a graphical method of Skeletal Shock Graph, which is based on the shape or geometry of the object. Shock Graph is an abstraction of the skeleton of a shape onto a Directed Acyclic Graph (DAG) in which the skeleton points are labeled according to the local variation of the radius function at each point. A large image database is created by using suitable image acquisition technique, which is converted into binary images. Skeleton and its labeling of the binary image is obtained by applying Skeletonization Algorithm. Then next steps adopted are formation of Shock Graph and labeled tree, indexing the database and generation of attribute vectors, pruning the data base and lastly matching the tree of query image with that of database images for recognition. This paper takes a critical review of the earlier contributions in this field and discusses the existing challenges and prospective research areas in Skeletal Shock Graph based object recognition.

\textbf{Keywords:} Directed Acyclic Graph, Skeletonization, Pruning of database, Skeletal Shock Graph, labeled tree.

\section*{I. Introduction}
One of the most fascinating features of the visual system of human being is “ability to quickly recognize a stimulus from a large visual memory”. Methods of object recognition and matching are broadly classified as intensity based methods relying on color, texture and geometry based methods, which use shape and structure of the object to match and retrieve from the query images. Most commonly used techniques of managing databases of image are Keywords or text-based technique, which is a common technique to provide information about the contents. It is suitable for texture based image with suitable segmentation. But for satisfactory degree of correctness & details, it requires large & complicated keyword system and manual annotations which are highly time-consuming, costly and dependant on the subjectivity.

The second technique is content based technique which involves generation of image description by analyzing content of image. The process is very common and simple. First, it retrieve image from database similar to the query image, this description store and extract features from query image and finally it estimates the similarity measures resulting image retrieval.

The third method is Graphical or Graph based method which is based on the shape or geometry of the object and uses a basic fact that humans show a remarkable ability to recognize objects without effort, despite the fact that objects may vary in color, texture, or size. For recognition purposes, objects are often represented as attributed graphs whose nodes represent their features (or their abstractions) and whose edges represent relations (or constraints) between the features. These graph representations allow us to express many perceptually significant object properties, such as geometric or hierarchical part structures.

Skeletal Shock Graph is an effective tool for object recognition, based on its geometry. The concept of Shock Graphs was first launched by Siddiqi and Kimia. Basically, Shock Graph is an abstraction of
skeleton of a shape onto a Directed Acyclic Graph (DAG) in which the skeleton points are labeled according to the local variation of the radius function at each point. Since its inception and evolution, it has been effectively used as a significant tool in the area of object recognition, which finds tremendous applications in automated or computer assisted environments.

Shape recognition is one of the primary goals of computer vision, allowing an image signal to be semantically labeled according to a priori knowledge of objects in the world. Early shape recognition work in the 60's and 70's focused on the categorization or generic object recognition problem, in which exemplar objects, i.e., specific object instances, were matched to coarse, prototypical models designed to be invariant to within-class shape and appearance deformation. Inspired by Blum's original idea of using directed graphs to define equivalence classes of shapes, Siddiqi and Kimia [1] defined the concept of a shock graph.

II. SKELETON AND SKELETONIZATION

Abstraction of object information is an important need for any automated or computer supervised or assisted environment for shape representation and modeling. Of course, the level of abstraction depends on the task for which the shape information is being required. Skeleton or medial axes have been extensively used for characterizing objects satisfactorily using structures that are composed of thin line or arc like patterns. Skeletonization is an image processing operation which reduces input shapes to axial stick-like representations. It has many applications like Industrial inspection, Expert system for Animal census, Intruder detection, Expert Security systems, Image / document processing, Biological part matching, Computer graphics, Robotics etc.

2.1 Medial Axis Transform

The medial axis transform (MAT) is basically described by using a medial axis and a radius function. The skeleton of a shape is like a stick diagram. One of the first formal definitions of the skeleton was proposed by Blum, who defined the medial axis of a shape as the locus of center points of all the maximal circles contained within the shape boundary. Blum called the shape's skeleton as the Medial Axis Transform (MAT). [28, 29, 30]

Blum also proposed an alternate definition of MAT in terms of a grassfire front-

The idea is to imagine that the shape is uniformly covered with grass and that fire is set simultaneously to all boundary points. The fire front will start propagating toward the center and will extinguish when meeting an opposing front. The collision points of the fronts form the medial axis of the shape. The distance of any skeleton point to its closest boundary points defines the time of formation of the points, and it corresponds to the radius of the maximal inscribed circle.

Skeleton is a representation of a 2D object which preserves the details of the silhouette of an object. Skeletons are converted to shock graph. Shock graph is nothing but shape abstraction which groups skeleton points according to the local variation of a radius function. These groups are called as Shock groups and are labeled as 1, 2, 3 or 4. Such a grouping will provide a decomposition of a skeleton into parts, whose interrelation will confirm to a well-defined grammar [1, 2, 31]. The shock graph is a directed acyclic graph and is obtained from these Shock groups which are the constituent primitives of the Shock Graph. The arrangement of shock graph is according to the Shock Graph grammar. Shock Graph is generated for each 2D image in the data base and is compared with the shock graph of the query image very efficiently and leads to an efficient technique for pattern matching.

Type 1 shocks form a segment of skeleton points, in which the radius function varies monotonically, as is the case for a protrusion. A type 2 arises at a neck, and is immediately followed by two type-1 branches bowing away from it in opposite directions. Type 3 shocks belong to an interval of skeleton
points, in which the radius function is constant. Finally, a type 4 arises when the radius function achieves a strict local maximum i.e. the boundary collapses to a single point.

![Types of shock graphs](image1.png)  ![Shock graphs based matching](image2.png)

**III. STEPS IN CONSTRUCTION OF SHOCK GRAPH**

The shock graph construction involves following steps:

**3.1. Labeling**

- \( R(s) \) be the radius function \( R : S \to \mathbb{R}^+ \)
- \( L(s) \) be a shock labeling function \( L : S \to \{1, 2, 3, 4\} \)
- \( S \) is the set of continuous medial axis points (i.e., shocks) of shape \( X \)
- Then \( L(s) \) for \( s \in S \) is defined as:

\[
L(s) = \begin{cases} 
4 & \text{if } \exists t > 0 \text{ s.t. } R(t) - R(s) \forall s \in N(s, t), \\
3 & \text{if } \exists t > 0 \text{ s.t. } R(t) - R(s) \forall s \in N(s, t) \neq \emptyset, \\
2 & \text{if } \exists t > 0 \text{ s.t. } R(t) < R(s) \forall s \in N(s, t) \text{ and } N(s, t) \neq \emptyset, \\
1 & \text{otherwise}
\end{cases}
\]

**3.2. Grouping**

Let \( B_1, \ldots, B_n \) be the longest groups of connected shocks in \( S \) s.t. \( \forall s, s' \in B_i, \ L(s) = L(s') \) and \( \forall s \in B_i, \) either \( |N(s)| \leq 2 \) or \( |N(s)| > 2, \) then \( s \) must be a terminal point of \( B_i \) (i.e., \( B_i \setminus \{s\} \) is connected), for \( 1 \leq i \leq n \). Let the group’s label, \( l(B_i) \), be the label of the label of shocks in \( B_i \), and similarly, let the time of formation of the group, \( t(B_i) \), be the interval \([\min(R(s)), \max(R(s))]\) defined by the time of formation \( R(s) \) for all \( s \in B_i \).

**3.3. Shock Graph**

The shock graph of a 2-D shape is a labeled graph \( G = (V, E, \gamma) \) such that:

- Vertices \( V = \{0, \ldots, n\} \), corresponding to the groups \( B_1, \ldots, B_n \) and 0 denoting a root node;
- Edges \( (i, j) \in E \subseteq V \times V \) directed from vertex \( i \) to vertex \( j \) if and only if \( i \neq j \) and \( i \neq 0 \wedge t(B_i) \geq t(B_j) \wedge B_i \cup B_j \) is a connected set of shocks, or \( i = 0 \wedge k \neq 0 \wedge (k, j) \in E \);
- Labels \( \gamma : V \to \{#, 1, 2, 4\} \) s.t. \( \gamma(i) = # \) if \( i = 0 \) and \( \gamma = l(B_i) \) otherwise.

The critical survey of related work indicates various steps involved in object recognition using shock graph are:

- Creating a large image data base by using suitable image acquisition technique.
- Converting an image of object or a 2-D shape into a binary image.
- Obtaining the skeletons of the binary images (Skeletonization Algorithm).
- Labeling the skeleton.
- Construction of the shock graphs and labeled tree from the skeleton.
- Indexing the data base and generation of attribute vectors.
- Pruning the database (if required) by comparing the shock graphs of objects in the database with the shock graph of the query image and retaining only the closest ones.
- Matching the tree of query image with that of database images for recognition.
Shock Graph offers the advantages that the labeling of skeleton points or shocks is invariant to translation and rotation of the shape and since the labeling is locally determined, it is stable under uniform scaling and within-class part deformation. Shock graphs offer a powerful and challenging framework for representing and matching qualitative shape [3]. But, quite limited efforts have been devoted to their use in realistic scenes in which a silhouette cannot be properly segmented. A careful examination of shock graph based recognition proves that most of the approaches and methodologies are conventionally applied to un-occluded, pre-segmented counter. Also, the overview of the research work in this field indicates that there is a tremendous scope in the area of shape recognition especially about recognition issues under the condition of image obscured and deformation. Hence it is necessary to focus on these issues and try to evolve a robust recognition based strategy, based on shock graphs.

IV. RELATED WORK

Although there has been quite extensive work in the area of shape representation and matching, shape recognition is still an open research issue [2]. With the time, many shape-matching approaches have emerged, but high space, time complexity and moderate recognition rate are still the limiting factors for their widespread acceptance as well as popularity. Sebastian et al. [4] use the full classification of the shock graph (MA endowed with a finer classification based on dynamics of flow) [5].

The shape recognition technique has to be robust to visual transformations like articulation and deformation of parts, viewpoint variation and occlusion. Thus, the shape representation has to effectively capture the variations in shape due to these transformations. In previous recognition applications, shapes have been represented as curves [6, 7], point sets or feature sets, and by medial axis [8, 9, 10, 11, 12] among others. 2D shape matching is achieved using hierarchical tree structure extracted from shock graph that in turn extracted from the skeleton of the shape of interest. It attempts to decompose a shape into a set of parts. Xiang Bai et.al.[15]matched skeleton graphs by comparing the geodesic paths between skeleton endpoints. In contrast to typical tree or graph matching methods, it does not consider the topological graph structure. The proposed comparison of geodesic paths between endpoints of skeleton graphs yields correct matching results. This method is able to produce correct results in the presence of articulations, stretching, and contour deformations. But the performance of this method is limited in the presence of large protrusions, since they require skipping a large number of skeleton endpoints.

Shahbudin S.et.al.[16] highlights the efficacy of the simplified shock graph (SSG) derived as a result of pruning from the original shock graph. It tests the feature vectors of the SSG with three different multi-class Support Vector Machines (SVM) classifier models namely the one-against-all (OAA), one-against-one (OAO) and the bias support vector machines (BSVM). As such, the validity of SSG as feature vectors to represent human postures is confirmed based on the findings from all three methods of SVM applied.

Macrini et.al. [17] presented an algorithm for identifying and representing the ligature structure, and restoring the non-ligature structures that remain. This leads to a bone graph, a new medial shape abstraction that captures a more intuitive notion of an object’s allies parts than a skeleton or a shock graph. It offers improved stability and within-class deformation invariance. Zaboli et. al. [19] presented a modified and optimized shock graph based method by adding branch points as key points to the shock graph and its grammar and consequently it developed a new grammar. But the recognition rate is not much attractive. Also, it had a limited focus on stability issues of the shock graph.

Behnam Hosseini et.al. [27] proposed some critical points on skeleton of silhouette are used for recognition and classification of human body movement and uses MLP with back-propagation for recognition of movement. But it does not focus the stability issues under articulation and deformation. Wei Zhao et.al. [34] proposed the use of skeletonization algorithm for removal of unwanted edges in processing of MRI images of skull and the brain tissue. It used the thresholding methods for this and use of skeletonization is limited to this purpose only.

Aaron D. Ward et.al. [25] proposed the G-MAT( Group-wise Medial Axis Transform), which is a group-wise skeletonization framework that yields a fuzzy significance measure for each branch, derived from information provided by the group of shapes for the determination of the pruning order.
of skeleton branches, using geometric and topological branch feature information provided by the group as a whole. Although the geometric and topological branch features used in this paper provided superior performance, there remain substantial opportunities for improvement using more sophisticated, pose-invariant features.

Jianhao Ding et.al. [13] discussed skeletonization method for the skeleton feature points extracted from human body based on silhouette images, which uses gradient of distance transform to detect critical points inside the foreground. Then, it converge and simplify critical points in order to generate the most important and elegant skeleton feature points and finally, presents an algorithm which connects the skeleton feature points and estimates the position of skeleton joints. But the results are unsatisfied and even encounter some mistakes when the joints are hidden or kept out. That is because this approach uses the uniform scales of the human body parts to determine the position of the inner joints such as elbow and knee joints, there might cause some errors between the scales of the different bodies.

Behnam Hosseini et.al. [14], presented skeleton of silhouette, being used for recognition and classification of human body movement and MLP with back-propagation is being used for recognition of movement. It has resulted average recognition rate of 92 %. But, it did not focus the stability issues under articulation and deformation.

An object recognition by comparing the sub-graphs by inserting a node in Shock branch whenever there is a sharp direction change has also been proposed by S. P. Hingway et.al. [32]. Addition of node improves the object recognition results. This changes the Shock graph topology. The object identification strategy called as Sub Graph matching method is used for Object recognition. But the improvement in result is not significant when the Skeleton does not have many curved segments.

V. LIMITATIONS AND SCOPE

Thus the critical review of the previous work gives rise to many unaddressed challenging research issues that need to be tackled effectively for the efficient object recognition. Scale and viewpoint changes may often originate new branches in a skeleton due to small details on the boundary that were not identified at a lower scale or that were not visible from a different viewpoint. These new branches can affect, to a certain extent, the structure a shock graph.

The shock labels depend on the pair-wise difference of the radius function at adjacent shocks which requires, in practice, to define some threshold to increase the tolerance to differences in radii.

Concave corners in the shape boundary may produce skeleton segments, where all the points are related only to the corner points at the boundary. Thus, small changes in the positions of these corner points will cause rather big changes in the ligature segments, which will ultimately give rise to differences in the corresponding shock graph. Noise may also have effect on the structure of shock graph.

The medial axis is sensitive to boundary perturbations (Boundary Noise), which introduce spurious edges in the graph. This may be typically tackled by regularization during the detection process or post-detection. Here regularization is an integral part of the recognition process.

Articulation and Deformation of Parts is also the issue of focus. The shock graph inherently segments a shape into parts and captures the hierarchical relationship between them. Hence it is robust to changes which may occur in some of the parts.

The viewpoint variations changes gradually, the spatial location and the shape of parts and it may be handled by deform edit. At certain views, there is a sudden appearance or disappearance of a part, which may be handled by the remaining edits. Appearance of a part may be handled by the splice edit. Shadows and highlights often lead to segmentation errors in figure-ground segregation. Highlights tend to cause small changes in the boundary, but may typically affect the shock-graph topology. Shadows are often more global in nature, but tend not to affect the shock-graph topology. Increasing Partial occlusion i.e. occluded blends with the shape or occluded blends with the background increases Shock Graph edit cost.

Hence, study of stability issues of Shock Graphs under various challenging situations may be the matter of focus. To accommodate this, robust method for labeling the skeleton points that will also accommodate small differences in scale may be one of the steps. Shock graphs define meaningful equivalence classes for shapes i.e. two shapes abstracted by the same graph are expected to be
perceptually similar at a coarse level. This property will allow us to index a model database to recover a set of candidate models with similar shock graphs. The next step may then be to verify each of these candidates by evaluating the information that we have ignored in the abstraction. With this goal in mind, a list of attribute vectors is associated with every point \( s \) in every node in the graph. Such vectors are 4-tuples \((x, y, t, \alpha)\) representing the attributes of each shock point in the node: Euclidean coordinates \((x, y)\), time of formation \(t\), and direction \(\alpha\). The set of attributes of each node may be used to define a local measure of similarity between two nodes in different graphs. Use of suitable AI technique like Fuzzy Logic, Neural Network or Hybrid tools for attribute vectors associated with every point may result into effective object recognition.

VI. EXPERIMENTAL RESULTS

![Figure 4](image1.png) **Figure 4**: Snapshot of the MATLAB window showing query image

![Figure 5](image2.png) **Figure 5**: Snapshot showing skeleton & labeling for query image

![Figure 6](image3.png) **Figure 6**: Snapshot of the MATLAB window showing shock-tree of query image.

![Figure 7](image4.png) **Figure 7**: Snapshot showing the result of tree matching with database for query image.

![Figure 8](image5.png) **Figure 8**: Snapshot of the complete GUI.
VII. IMPLICATIONS

The proposed research outcome will definitely be a valuable contribution in the vast and ever open field of shape recognition. The outcome of the proposed research work may have several applications in the fields of Automated Industrial inspection, Expert system for Animal census, Intruder detection, Expert Security systems, Image / document processing, Biological part matching, Computer graphics, Robotics etc. The labeling of skeleton points or shocks is invariant to translation and rotation of the shape. Since the labeling is locally determined, it is stable under uniform scaling and within-class part deformation. But all these issues are always posed to the challenging situations like boundary noise, articulation and deformation of parts, view point variations and partial occlusion. So efforts will be concentrated on improving the stability of shock graph for object recognition under these challenging issues. Further, it is expected that the use of suitable AI technique will make the recognition process fast and accurate.

VIII. CONCLUSION

Skeletal Shock Graph is an effective tool for object recognition, based on its geometry Object recognition is one of the primary goals of computer vision, allowing an image signal to be semantically labeled according to a priori knowledge of objects in the world. Scale and viewpoint changes, concave corners in the shape boundary, boundary perturbations, viewpoint variations, segmentation errors due to shadows and highlights, increasing partial occlusion are still some of the challenging situations in object recognition, may be the matter of focus for the researchers. To tackle this, robust method for labeling the skeleton points that will also accommodate small differences in scale may be one of the steps. Shock graphs define meaningful equivalence classes for shapes which allow us to index a model database to recover a set of candidate models with similar shock graphs. Thus, an improved skeletonization, improved indexer, improved matcher and use of suitable AI tool may further enhance the recognition efficiency of the system, especially under the challenging situations for a large images data base of varieties of shapes concerned for varieties of object recognition applications in automated or computer assisted environments.

REFERENCES


Authors

Pankaj M. Pandit is working as Associate Professor in Dept. of Electronics & Tele. Engg. of Jawaharlal Darda Institute of Engg. & Technology, Yavatmal (M.S.) He completed B.E. in Electronics & Telecomm. Engg. from Govt. College of Engg., Amravati in 1994 and then M.E. with specialization of Digital Electronics from P.R.M.I.T.& R., Badnera. He also completed M.B.A. with marketing as specialization from YCMOU, Nashik in 2007. Currently he is pursuing Ph.D. in Amravati University. He has 15 years of teaching experience and contributed 8 research publications at National and International Journals and Conferences. His areas of expertise and interest are Analog & Digital Electronics, Artificial Intelligence, Digital Image Processing, Machine Intelligence, Computer Vision, Entrepreneurship Development, Consumer Behavior, Lifestyle Marketing, and Consumer Psychology. He is Life Member of I.S.T.E. and IETE. He is also contributing as Executive Member of IETE Center, Amravati and ISF Faculty Advisor of the present Institution.

Sudhir G. Akojwar, is working as Professor of Electronics and Tele-Communication Engineering and Principal at Dr Bhausaheb Nandurkar College of Engineering & Technology, Yavatmal(M.S.). He had 22 Years of teaching experience and published 27 research publications at National and International conferences and journals. He completed Ph.D. from VNIT, Nagpur. His area of research is WSN, Image processing, Neural Networks and Embedded Systems. He completed his B.E.in Instrumentation Engineering form Shri. Guru Gobindsingh College of Engg. and Tech., Nanded, in 1990. He joined as lecturer in the Department of Electronics Engg. at RCERT in 1990. He completed his M.Tech. in Electronic Instrumentation from REC, Warangal [presently NIT, Wrangal]. He was promoted as Asst. Prof. and then as Professor in the same college. He worked as Secretary and now as a Chairman of the local ISTE-Chapter. He worked as Chairman of the board of studies in Information Tech. at R.T.M. Nagpur university. He is member of IEEE and IEEE computer society. He is also member of ISTE, IETE, IE and BMESI.