**Remittances Review**
January, 2023
Volume: 8, No: 1, pp. 434-452
ISSN: 2059-6588 (Print) | ISSN 2059-6596 (Online)
remittancesreview.com

Received: 11 November 2022 Accepted: 15 March, 2023
DOI: https://doi.org/10.33182/rr.v8i1.28

# A Study on the Visual Rhetoric of Luoshan Shadow and its Symbolic Representation Based on Improved VSLAM Algorithm

Taotao Xu [*,1,2],

***Abstract***

*In this paper, we first study the implementation process of the improved VSLAM algorithm. After initializing the root node containing all the features, a quadtree node partitioning is performed on the image and the network grid of each layer of the ginormous tower. And FAST feature points are extracted by adaptive thresholding within the grid and a specific number of feature points are saved according to the size of Harris response values within the grid. Then the algorithm is experimentally tested for feature point distribution uniformity, matching performance and SLAM performance, and compared with the false matching coarse screening algorithm. Finally, the improved VSLAM algorithm is applied to extract and analyze the symbolic representations, image rhetoric and visual segment rhetoric of Luoshan shadow. In terms of representation matching, the VSLAM algorithm matches at 50%-60%, and the improved matching algorithm matches at 70%-80% with an accuracy rate of 20%. In terms of image rhetoric, 87 shadow works use anthropomorphic rhetoric, the number of works using repetition, pun, reference and exaggeration rhetoric is about 100, and the largest percentage of examples is about 220, and 100 shadow works use accumulation and comparison rhetoric. The improved VSLAM algorithm can be good for the extraction of Luoshan shadow representations and rhetoric, which is beneficial to the study of shadow.*

***Keywords:*** *VSLAM algorithm, Luoshan shadow, visual rhetoric, symbolic representation, feature extraction.*

## Introduction

Luoshan shadow puppet is an excellent folk art with a long history in southern Henan, and is a comprehensive folk art that concentrates folk opera, folk musical instruments, folk art and folk craft (Binyon, 1966; Mendiola A ). Among the historical southern Henan culture, Luoshan shadow puppets, as one of the most representative outstanding folk cultures in Xinyang, Henan Province, bring visual beauty with their rich and unique shape and color features, and are unique in the shadow culture (Yan C 2016). Luoshan shadow play has distinctive characteristics of local theater art of Luoshan (Qin H W 2015). The literature (L, 2019) examines the classification of shadow modelling and concludes that there is consistency in the modelling style of early shadows. The literature(E, 2016) summarizes the four artistic characteristics of shadow modeling: reasonable deformation, bold exaggeration, clever symbolism, and elegant and simple decoration. The

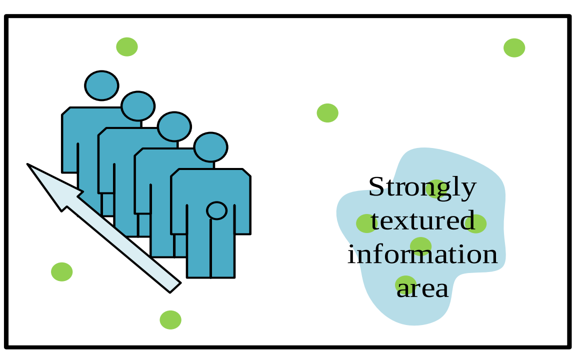
[1] Faculty of Applied and Creative Arts, Universiti Malaysia Sarawak, Kota Samarahan, 94300, Malaysia
**Corresponding author: Wei Zhang** (unimas2021@126.com)
[2] College of Art and Design, Xinyang University, Xinyang, Henan, 464000, China

literature(A, 2016) introduced the different shapes of shadow with a lot of vivid pictures and text illustrations, and made a more detailed regional division of the shadow shapes, and introduced and analyzed their different stylistic features. The literature(Q, 2018) introduces the characteristics of the origin, shadow carving, shadow tunes, sung roles, and circulating repertoire of the shadow play in Luoshan, and has some numerical statistics on the current status of the shadow play in Luoshan. Literature(Q. H, 2018) explored the historical mechanism of the inheritance and development of Luoshan shadow play in terms of the basis of devotional beliefs, the way of memory transmission, the listening habits of the people, and the homogeneous vernacular environment. The so-called "internal knowledge" of the script refers to the knowledge of the "play path" of Luoshan shadow play, which is the outline of the script and cannot be performed directly. The literature (M, 2015)points out that the singing style of Luoshan shadow is divided into five categories: Sheng, Dan, Jing, Mao, and Ugly, and there is a distinction between Eastern and Western tunes, and gives a brief introduction to the art of Luoshan shadow as a whole(L. J. H, 2018). The paper (C. Y, 2016)briefly analyzed the development history and current situation of Luoshan shadow puppet theater in Henan Province, and expended that under the environment of modern design, Luoshan shadow art should seize the opportunity and accelerate the reform of its own expression, while the local government should increase the support and investment to protect the contemporary Luoshan shadow art. The literature (Y, 2017),starting from the interaction between dramatic performing arts and folksy rituals, shows that shadow puppet theater and folksy rituals are an inseparable whole through the analysis of the meaning of repayment rituals and folksy life rituals. The literature (Zhang J 2017)summarizes the question of how Luoshan shadow play was formed and developed from three aspects: historical development, the origin of Luoshan shadow play, and Luoshan shadow play after liberation. The literature (S. Y, 2016)mentions the historical origins of Luoshan shadow play, and also mentions that from the founding of New China to the beginning of the 21st century, Luoshan shadow play characters, singing design and musical instrument accompaniment were influenced by the political and economic changes in different periods. The literature (Özcan, 2002) mentions the rapid economic development and the improvement of people's living standards since the reform and opening up, bringing both benefits and disadvantages to the development of Luoshan shadow puppets, presenting the current situation of the development of Luoshan shadow puppets. The literature (Currell, 2015) briefly describes the history and development status of Luoshan Piyin, and proposes that under the environment of modern design, Luoshan Piyin should seize the opportunity and accelerate the reform of its own expression, while the local government should increase support and investment to protect the contemporary Luoshan Piyin. The paper (Bryant, Heard, Watson, & MacDonald, 2002) analyzes and summarizes three characteristics of the shadow form, harmony and rhythm, the rude and ancient color of the shadow, and the fine and elaborate craftsmanship of the shadow from the five production processes of making the skin, overdrafting, engraving, coloring, fixing and embellishing. The paper (S, 2019)discusses how to use dots, lines and surfaces to show the visible two-dimensional image in two-dimensional space, how to express the personality characteristics of shadow characters through allegorical patterns and colors, and how to use colors in the process of artistic creation, thus presenting the characteristics of the Luoshan shadow in terms of shape and colors. This paper firstly explores the implementation process of the improved VSLAM algorithm by meshing each layer of the image pyramid network, then extracting FAST feature points by adaptive thresholding within the mesh, and saving a specific number of feature points according to the size of the Harris response value within the mesh. The algorithm improvement is proposed for the VSLAM algorithm features are too uniform, resulting in retaining a large number of poor quality features, and the algorithm is experimentally tested for feature point distribution uniformity, matching performance and SLAM performance, and the improved VSLAM algorithm and VSLAM algorithm are allowed to be compared. Finally, the algorithm is applied to extract and analyze the symbolic features and the image rhetoric and visual segment rhetoric used in Luoshan shadow.

## Improve VSLAM algorithm

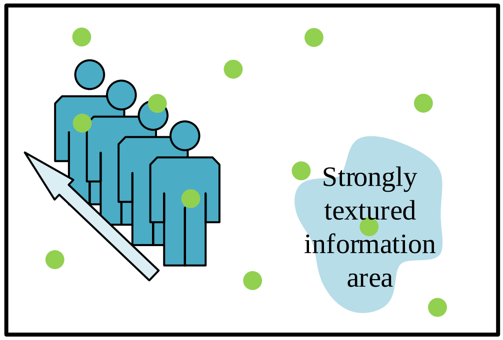
### Analysis of standard VSLAM algorithm

VSLAM feature points have been widely used in SLAM systems because of their better real-time performance and accuracy, but there are still some potential problems. As shown in Figure 1, the feature extraction results of the standard VSLAM algorithm in a complex scene, which contains high-speed moving pedestrians as well as a region with strong texture information. The high dynamic motion pedestrian leads to motion blur in the region where it is located, and the small grayscale difference between pixel points in this region leads to the poor quality of the features extracted on it, which is not good for tracking. However, since the feature retention strategy of the VSLAM algorithm is based on the Harris response value, which is a measure of the grayscale difference between a feature and its surrounding pixels, the larger the grayscale difference, the higher the response value. Therefore, most of the features extracted by the standard VSLAM algorithm are concentrated in regions with strong texture information, and these local features do not reflect the global information well and introduce errors for the SLAM system. In view of this, it is particularly important to study a VSLAM feature point homogenization algorithm.




(a) Standard ORB algorithm

**Remittances Review**
January, 2023
Volume: 8, No: 1, pp. 434-452
ISSN: 2059-6588 (Print) | ISSN 2059-6596 (Online)




(b) Quadtree-based ORB algorithm

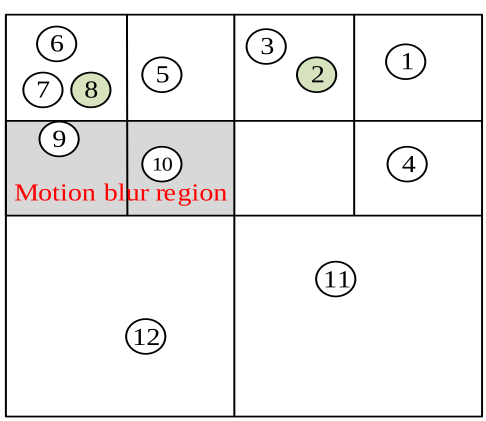
**Figure 1.** VSLAM feature extraction example in complex scenes

## Improved Qtree_VSLAM feature point homogenization algorithm

A Qtree_VSLAM algorithm is used to extract features in VSLAM-SLAM2 to avoid feature aggregation. However, some potential problems are ignored. Firstly, the standard Qtree_VSLAM algorithm defects are discussed, and then an improvement strategy is proposed to enhance the standard algorithm.

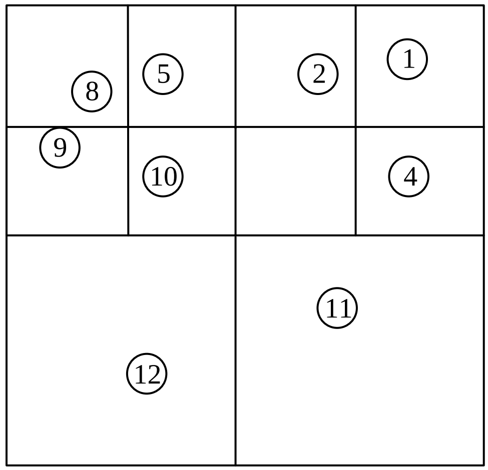
## Analysis of Qtree_VSLAM feature point extraction algorithm

Unlike the standard VSLAM algorithm which is easy to aggregate, the Qtree_VSLAM algorithm extracts VSLAM features uniformly but tends to retain some low quality features. A schematic diagram of the quadtree-based VSLAM feature point extraction algorithm is shown in Figure 2. The image with 12 VSLAM features is first partitioned into quadtree nodes, and then the features in the nodes are retained according to their Harris response values. Figure 2(b) shows the final VSLAM feature points retained by the Qtree_VSLAM algorithm with the following detailed feature point retention strategy:

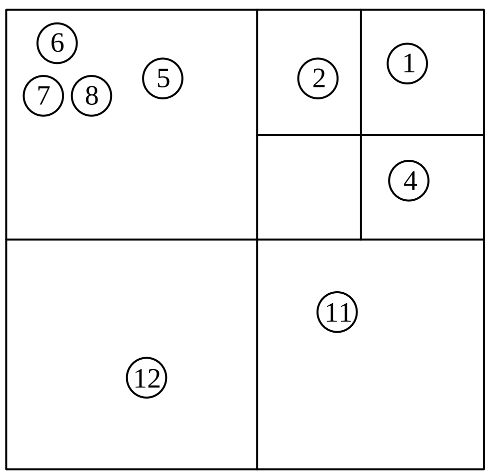



(a) Quadtree-based segmentation

January, 2023
Volume: 8, No: 1, pp. 434-452
ISSN: 2059-6588 (Print) | ISSN 2059-6596 (Online)




(b) stand Quadtree-based algorithm




(c) improved Quadtree algorithm

**Figure 2.** The feature extraction example of Quadtree-based VSLAM algorithm

Step 1: Initialize the root node which contains all the features and the root node will be used as the parent node for the next partition.

Step 2: Split all parent nodes into quadtree nodes and remove the child nodes without features.

Step 3: If the number of child nodes containing only one feature is greater than the number of desired feature points, move to step 4. otherwise, move to step 2.

Step 4: Keep the feature point with the highest Harris response value in each child node as the final VSLAM feature point.

Volume: 8, No: 1, pp. 434-452
ISSN: 2059-6588 (Print) | ISSN 2059-6596 (Online)

Step 3 is the quadtree node segmentation exit condition, which determines the segmentation depth and feature uniformity. However, in a highly dynamic environment, this node segmentation method may lead to some problems. The gray part indicates the motion blur region, which has little gray level difference between pixels. The features extracted in the motion blur region have a low Harris response value, which is not favorable for subsequent matching and tracking. The results of node segmentation and feature point retention when setting the required number of features to 9 are shown in Fig. 2(a)(b), respectively. As shown in Fig. 2(b), the poor quality features 9 and 10 are retained, which is caused by over-segmentation of the quadtree nodes. These poor quality features are then fed into the visual odometer, which leads to errors in the SLAM system. The above example is only caused by the presence of moving objects in the scene, and the inter-pixel grayscale difference is also small in regions with weak texture information (e.g., screen, floor, table, etc.), where poor quality features are also retained for the same reason. In each segmentation of the parent node, it is necessary to re-determine to which child node the features in the parent node belong. Therefore excessive segmentation of quadtree nodes can seriously affect the efficiency of the algorithm operation.

Based on the above analysis, it can be concluded that there are two main defects in the Qtree_VSLAM algorithm applied to the SLAM system in dynamic scenes: the feature points are distributed too evenly, so that a large number of poor quality features are retained, which affects the localization accuracy of the SLAM system. Secondly, the quadtree nodes are divided too much, which affects the operation efficiency of the system.

**Improved Qtree_VSLAM feature point extraction algorithm**

Image pyramids can express multi-scale information by constructing multiple resolutions, which can well solve the scale problem in images. To reasonably distribute the distribution of features to be extracted, the number of features extracted in each layer of the image pyramid is set as follows:

$$DesF_i = N \times \frac{1 - InvSF}{1 - InvSF^n} \times InvSF^i \tag{1}$$

Where, $N$ is the total number of desired feature points, $n$ represents the total number of pyramid layers, $DesF_i$ is the total number of features to be extracted from the $i$ th pyramid layer, and $InvSF$ is the inverse of the scale factor.

When the image ginta VSLAM algorithm. In VSLAM-SLAM2, the FAST feature extraction thresholds are set based on empirical engineering values and do not adapt well to the external environment. For example, high thresholds cannot extract enough features in image regions with weak texture information. In view of this, the extraction thresholds are adaptively processed by setting the initial threshold $iniTH$ as:

$$iniTh = \frac{1}{\kappa \cdot ni} \sum_{x=1}^{ni} (I(x) - \kappa)^2 \quad (2)$$

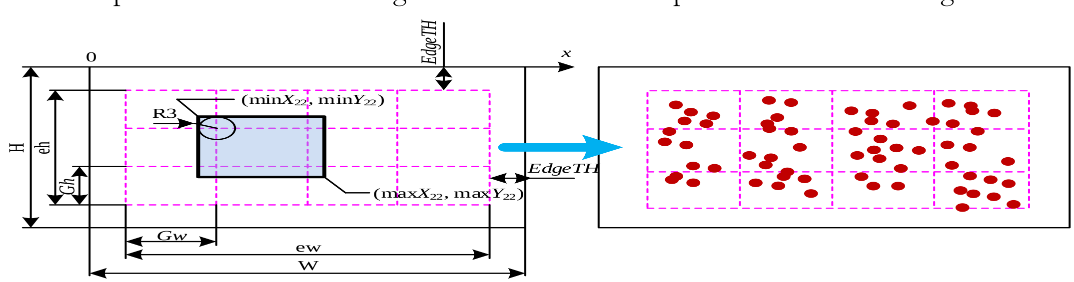
Where $I(x)$ represents the $x$ nd pixel gray value, $\kappa$ is the average of the gray value of each pixel on the image, and $ni$ represents the total number of pixels. To divide the image into quadtree nodes, unlike the standard algorithm, the improved algorithm limits the depth of quadtree node division. The relationship between the number of quadtree nodes and the depth of node division can be expressed as:

$$CNodes = 4^d \quad (3)$$

The feature retention strategy in quadtree nodes is the most challenging task to improve the algorithm. Figure 2(c) shows the feature extraction schematic of the improved algorithm, where the segmentation iteration depth is set to 2. It can be seen that the improved algorithm no longer retains dynamic features because features 6 and 7 have higher Harris response values and are higher than the features in the motion blur region, and the improved algorithm tends to retain features with higher Harris response values in the nodes, effectively filtering out low-quality features. The algorithm describes the node division and feature retention strategy in detail. Where $\max D$ is the maximum depth of the iteration, $nNum(i)$ indicates the number of feature points in the $i$ rd node, the feature point with the maximum Harris response value in the node is denoted as $\max kp$, and $\min H$ is the minimum acceptable Harris response threshold.

**VSLAM feature point homogenization algorithm based on grid partitioning**

In this paper, we introduce a new grid-based VSLAM feature homogenization algorithm to extract VSLAM features uniformly. As Figure 3 shows the flow of Grid_VSLAM algorithm feature extraction. Firstly, the grid is divided for each layer of the image pyramid network. Then FAST feature points are extracted by adaptive thresholding within the grid, and finally, a specific number of feature points are saved according to the size of Harris response values within the grid.




( a) Meshing ( b) FAST feature extraction in the grid

**Figure 3.** Image feature point uniform extraction algorithm based on grid division

Volume: 8, No: 1, pp. 434-452
ISSN: 2059-6588 (Print) | ISSN 2059-6596 (Online)

**Mesh model determination**

Construct an image pyramid with layer $n$, and denote the image of layer $i$ of the image pyramid as $I_i$. The image of layer $I_i(i=1,2,\cdots,n)$ can be obtained by downsampling the image of layer $I_{i-1}$ of the pyramid by a factor of $SF$, and the scaling scale of each layer can be expressed as:

$$S_i = SF^i (i=0,1,\cdots,n-1) \quad (4)$$

Where, $n$ is the number of layers of the image pyramid, $SF$ is the scale factor of each layer of the image pyramid. $S_i$ is the scaling of the image at the $i$ th level. The scale increases sequentially with the increase of the number of layers of the 8-layer ginormous tower. The scale problem of VSLAM algorithm is effectively solved by constructing image gin towers to add scale information to images.

After obtaining the image pyramid with $n$-layer structure, each layer of the image pyramid is meshed. If the total number of VSLAM feature points to be extracted is $N$, the number of feature points to be extracted in each layer of the image pyramid is set to be:

$$DesiredF_i = N \times \frac{1-InvSF}{1-InvSF^n} \times InvSF^i \quad (5)$$

where $DesiredF_i(i=0,1,\cdots,n-1)$ is the number of feature points needed for layer $i$, $n$ represents the total number of layers of the image pyramid, and $InvSF$ is the reciprocal of the scale factor, which can be expressed as $InvSF = 1/SF$.

Considering that the image edge feature points are easily lost in two consecutive frames, which leads to matching failure, the boundary is set on each frame of the image. The boundary threshold of the image feature points is noted as $EdgeTH$, which indicates the distance from the extracted boundary of the image feature points to the image boundary. The boundary division of the image is shown in Figure 3(a). Denote the width $ew_i$ and height $eh_i$ of the extracted region of feature points of the $i$ nd layer image ginta, whose values can be expressed as the following equation:

$$ew_i = w_i - 2EdgeTh \quad (6)$$

$$eh_i = h_i - 2EdgeTh \quad (7)$$

where $w_i$ and $h_i$ denote the width and height of the image of the $i$ rd level of the image pyramid, respectively.

The area of each grid should be as equal as possible to make the feature points more evenly distributed. The number of rows $Rows_i$ and columns $Cols_i$ of the grid divided by image layer $i$ of the image pyramid can be obtained as:

$$\begin{cases} Cols_i = \sqrt{\dfrac{DesiredF_i}{ImRat_i}} / t \\ Rows_i = ImRat_i \cdot Cols \end{cases} \quad (8)$$

where $t$ is the mesh division factor, and the number of mesh divisions decreases when $t$ increases. $ImRat_i$ denotes the image aspect ratio of the $i$ th level of the pyramid, which is $ImageRatio_i = Rows_i / Cols_i$.

If the grid width and height of layer $i$ image are $Gw_i$ and $Gh_i$ respectively, then the number of rows of the grid, the width and height of the grid, and the number of grids can be expressed as:

$$\begin{aligned} Gh_i &= eh_i / Rows_i \\ Gw_i &= ew_i / Cols_i \end{aligned} \quad (9)$$

The extraction of FAST feature points requires the detection of pixel points on the circumference of a circle with radius 3 of the point to be extracted. Therefore, to prevent the grid boundary points from not being detected after the grid division, which leads to the loss of image information points, the edges of the grid are expanded. Each expanded grid can be determined by four boundary points $(\min X, \min Y)$, $(\min X, \max Y)$, $(\max X, \max Y)$ and $(\max X, \min Y)$, and each parameter of the four boundary points of the grid can be determined by the following equation:

$$\min X_{mn}(i) = n \times Gw_i - Th \quad (10)$$

$$\min Y_{mn}(i) = m \times Gh_i - Th \quad (11)$$

$$\max X_{mn}(i) = \min X_{mn} + Gw_i + Th \quad (12)$$

$$\max \mathrm{Y}_{mn}(i) = \min \mathrm{Y}_{mn} + Gh_i + Th \quad (13)$$

Where, $\min X_{mn}(i)$ and $\max X_{nn}(i)$ represent the minimum and maximum values of the horizontal coordinates of the $m$ th row and $n$ th column grid on the image of the $i$ rd level of the image pyramid, respectively. $\min Y_{mn}(i)$ and $\max Y_{mn}(i)$ represent the minimum and maximum values of the vertical coordinates of the grid in row $m$ and column $n$ on the image of the $i$ th level of the image pyramid, respectively. $Th$ is the expansion value of the grid.

As shown in Fig. 3(a), a schematic diagram of a grid division with width $W$ and height $H$ is shown, and the dotted lines enclose 12 grids in 3 rows and 4 columns, which represent the feature point extraction regions. The blue filled area in the figure is the expansion of the grid in row 2 and column 2 on the image, where the boundary points of the four expansion areas are derived from Eqs. (10)-(13). The circle of radius 3 in the figure is a schematic diagram of the detection of the feature points on the boundary of the region to be detected, and it can be seen that the feature points on the boundary of the grid can be detected smoothly after the expansion is performed.

**Intra-grid adaptive threshold feature point extraction**

When the grid of each layer of the image pyramid is determined, FAST feature points are extracted in each grid. In order to make the distribution of feature points in the grid of each layer of the pyramid as uniform as possible, the desired number of feature points extracted in each grid of the $i$ st layer of the pyramid is set to:

$$cDesF_i = \frac{DesF_i}{Rows_i \cdot Cols_i} \tag{14}$$

The feature extraction algorithms in the standard VSLAM algorithm and the VSLAM-SLAM2 algorithm are based on fixed engineering empirical thresholds to extract FAST feature points, which cannot adapt well to environmental changes. Therefore, the improved method uses an adaptive thresholding method in the grid to extract FAST feature points, and sets the initial threshold as follows:

$$iniTh = \frac{1}{\kappa \cdot ni} \sum_{x=1}^{ni} (I(x) - \kappa)^2 \tag{15}$$

where $I(x)$ is the gray value of each pixel in the image, $\kappa$ is the average gray value in the image, and the total number of pixel points in the image is denoted as $ni$. The FAST feature points in the grid are extracted with the initial threshold $iniTh$, and the number of extracted points is $iniN$. If $iniN$ is less than the required number of feature points $cDesF_i$ in the grid, set the minimum threshold $\min TH$ as $1/4$ of the initial threshold, and continue to extract the FAST feature points in the grid with the minimum threshold $\min TH$ to complete the adaptive extraction of feature points in the grid.

January, 2023
Volume: 8, No: 1, pp. 434-452
ISSN: 2059-6588 (Print) | ISSN 2059-6596 (Online)

After pre-extracting the feature points in the grid, most of the feature points in the grid are more than the expected number of feature points $cDesF_i$, among them, there are a small number of feature points in the grid are less than the expected number of feature points $cDesF_i$. How to filter the feature points needed for the layer network among the many feature points in each grid $DesF_i$, this paper adopts the following scheme, whose steps can be summarized as follows:

Step 1: Construct an image pyramid for the input image and calculate the expected number of feature points $DesF_i$ for each layer according to the scale factor of that layer.

Step 2: On the $i$ st layer of the pyramid, the number of grids is divided into grids of $nGrid_i$. FAST feature point extraction with adaptive thresholding is performed for each grid.

Step 3: Define the grid containing at least one feature point as a valid grid, and sort the feature points in the valid grid according to the size of Harries' response value from largest to smallest, and record the sequence as $OrderKp$ and its number as $ov$. Judge the size of the number of feature points in the sequence and the expected number of feature points in the grid as $cDesF_i$. If it is larger than $cDesF_i$, take the first $cDesF_i$ items in $OrderKp$ and deposit them into the temporary feature point sequence $tmpKp$. If it is less than $cDesF_i$, deposit all the feature points in $OrderKp$ into $tmpKp$. The feature points deposited into the $tmpKp$ sequence are eliminated from the grid they belong to in turn.

Step 4: Calculate the difference $DifNum_i$ between the expected number of feature points $DesiredF_i$ in layer $i$ of the image pyramid and the temporary feature point sequence $tmpKp$ if $DifNum_i > 0$, sort the remaining feature points in all the grids in that layer according to the Harries response value from largest to smallest, record the sequence as $OrderKp$, and keep the first $DifNum_i$ items to be stored in $tmpKp$. Otherwise, skip to step 5 directly.

Step 5: Store the temporary feature point sequence $mmpKp$ into the final feature point sequence $Kp$.

Step 6: If $i < 8$, set $i = i + 1$, jump back to step 2 again and execute in order. Otherwise, the algorithm terminates.

The flow of VSLAM feature point homogenization algorithm based on grid division is shown in Algorithm 2. Through the above steps, the process from image ginta construction to the division of each layer of the image ginta grid and the extraction of its adaptive feature points within the grid is completed, and then to the screening process of the grid feature points within the image ginta. The final feature point sequence $Kp$ is obtained, which is the final VSLAM feature points extracted by the algorithm.

**Experiments on visual representation extraction under shadow dataset**

In this section, 10 videos of Luoshan shadow shadow under the dataset are selected to extract continuous images, and five images are extracted from each video and divided into 10 groups. The VSLAM algorithm and the improved VSLAM algorithm are compared and analyzed, and the data are plotted for three indexes: the total number of characterization points, the total time spent in the feature extraction phase, and the average time spent in a single point.

The average elapsed time for feature extraction of the improved VSLAM algorithm is lower than the average elapsed time for feature extraction of the VSLAM algorithm. The reason for this is that in the process of calculating the descriptors, the calculation of the center of mass of the significance values in the neighborhood space of the feature points is increased, and the increased computation time of the above process is greater than the reduced computation time of eliminating the redundant regions. However, the original VSLAM algorithm takes about 22ms, and the improved VSLAM algorithm is between 15ms-20ms, and the performance of the improved algorithm is better than that of the VSLAM algorithm. In addition, the feature values extracted by the improved algorithm are basically around 1000-2500, while the feature extraction values of the VSLAM algorithm are around 500-1500, which shows that the improved VSLAM algorithm extracts more representations than the original algorithm. Figure 4 shows the comparison of algorithm visual representation extraction.

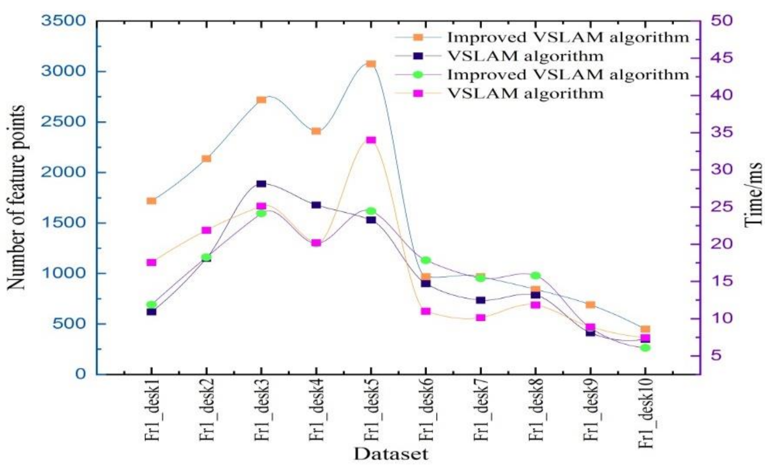



Figure 4 Comparison of visual representation extraction

Volume: 8, No: 1, pp. 434-452
ISSN: 2059-6588 (Print) | ISSN 2059-6596 (Online)

## Feature matching experiments under the shadow dataset

Comparing and analyzing the three algorithms of VSLAM algorithm, false match coarse screening improvement and the improved VSLAM algorithm in this paper, the matching accuracy index was calculated for the three indexes of total number of matches, correct number of matches and matching time consumption. The experiments were repeated five times for each group of images in the selected dataset, and the average value of the five repeated experiments was recorded for the total number of matches and the number of correct matches, and the matching accuracy rate was calculated at last.

The matching accuracy of the original VSLAM algorithm features was between 40%-50%. After improving the matching algorithm, the matching accuracy was between 70%-80% for matching between image 1 and image 2 where the image changes were not obvious, at 70% for the bike and grafl groups, at 50% for the boat group, and at 80% for the leuven and boat groups. In the matching of image 1 with image 3 where the image changes significantly, the matching accuracy of the grafl group is below 60%, the leuven group is below 70%. The matching accuracy of the bike group is 50% and the boat group is at 60%. The accuracy of the improved VSLAM algorithm is improved by about 20% in the matching accuracy of each group. Figure 5 shows the comparative analysis of the correct matching rate of the improved VSLAM.

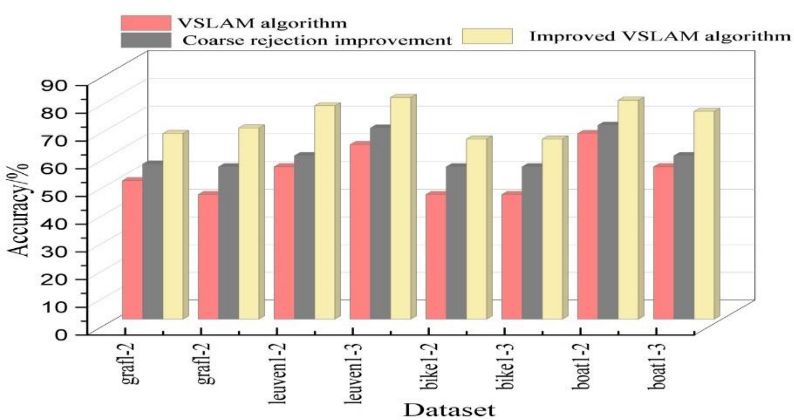



Figure 5 Comparative analysis of correct matching rate for improved VSLAM

## Symbolic representation and visual rhetoric of Luoshan shadow based on improved VSLAM algorithm

### Definition of symbols

Symbols encompass the signs themselves, such as signals and pointers, that we see, as well as the users of the signs and the things and meanings to which they refer. The scope of the study of signs is very broad, and therefore the definition of signs is also very broad. Saussure divides linguistic signs into their energetic and referential components, and calls the process of combining the energetic and referential into one the action of meaning. The action of meaning conveys a message clearly to the public and evokes a new meaning in people's minds. Saussure's definition put an end to the chaotic concept of the term "sign" in linguistics, and the study of the energy, reference, and meaning of signs became the focus of research on linguistic signs and even semiotics in general. Under the influence of Saussure's view of signs, Yermeslev proposed a hierarchy of signs, pointing out that the sign as a whole consists of an "expression layer" and a "content layer", which correspond to Saussure's "referent" and "referent" respectively. "These two levels are linked by relations, and the whole of the levels and their relations form a system of meanings. Table 1 shows Yermeslev's view of symbols.

**Table 1.** Yermeslev's view of symbols

| Symbols | Expression layer | | Content layer | |
|---|---|---|---|---|
| | Endoplasm | Form | Endoplasm | Form |
| | Representational elements | Composed of combination rules and aggregation rules | he referred, emotional, and conceptual, ideological, etc. | The form of the relationship between the referred organisational relationships on |

The expression of visual symbols should consist of visual representation elements, including lines, forms, colors and phonemes. The expression form is composed of combination rule and aggregation rule, the combination rule is that each representational element combines with each other to form a complete symbol form, which helps to highlight the theme, and the aggregation rule is that multiple elements are placed together to reflect each other, which helps to render the atmosphere.

The content of the content is the reference of the symbol, including the meaning, emotion, concept, consciousness and other aspects of the reference. The content form is the organizational relationship between the referents, and we consider the organizational relationship of the referents as rhetoric. Table 2 shows the symbolic analysis.

**Table 2.** Symbolic analysis

| Symbols | Expression layer | | Content layer | |
|---|---|---|---|---|
| Hand | Inner quality | Inner quality Form | Inner quality Form | Form |
| | Strong black and white large colour blocks | The elements are grouped with each other | Referring to this arm or the person is in a different state of being | The use of metaphors metaphorical expression |

January, 2023
Volume: 8, No: 1, pp. 434-452
ISSN: 2059-6588 (Print) | ISSN 2059-6596 (Online)

The meta-linguistic system allows us to explain how language is constructed through the shadow language. We apply the framework of the meta-linguistic system to a specific shadow plot. The expression layer of the Luoshan shadow symbol system is the withered arm that is being stretched, and the content layer is the actor's psychological distortion and difficulty in self-control.

---

This shadow symbolic text as a whole becomes the content layer of its title symbolic text, while the expression layer adopts a contrasting black and white color and a stiff and twisted arm image to express the key content of the overall film. We can see that the episodes shown in the film can express certain meanings, and in the credits, these episodes and meanings are re-copied and become energetic references of the shadow episodes and meanings by the symbolic language of the credits. Table 3 shows the symbolic system transformation.

**Table 3.** Symbol system conversions

| Symbol system | Expression level | Content layer |
|---|---|---|
| Title symbol system | Hand | Pain in the hands |
| Shadow symbol system | Covered hands | The characters are psychologically distorted and have difficulty controlling themselves |

**Image rhetoric of Luoshan shadow based on improved VSLAM algorithm**

Images are the basic constituent units of visual texts, and images will also constitute with each other by some rhetorical means of images-a visual discourse with a stable structure and complete meaning. Therefore, the exploration of the rhetoric of images in Luoshan shadow puppets should be carried out from these two levels. The rhetoric of pictorial images in Luoshan shadow puppet includes imitation, metaphor, borrowing, exaggeration, representation, anthropomorphism, symbolism, reference and other rhetorical patterns.

A statistical survey of the works of Luoshan shadow puppets shows that imitation is the most frequently used pictorial rhetoric, with about 500 works using imitation rhetoric. After that, symbolic rhetoric is used, with a frequency of about 300.

Then comes anthropomorphism, representation, and quotation, with frequencies ranging from 120 to 150 respectively.

The rhetorical expressions of metaphor, exaggeration, borrowing and symbolism, which are commonly used in Luoshan shadow puppets, are less frequently used. The rhetorical patterns of imitation, personification and quotation are typical of the rhetorical methods commonly used in narrative-written images, mostly reproducing the objective world, while metaphor, symbolism and exaggeration are often found in themes with strong ideological overtones such as religion and politics. Figure 6 shows the rhetorical statistics of images.

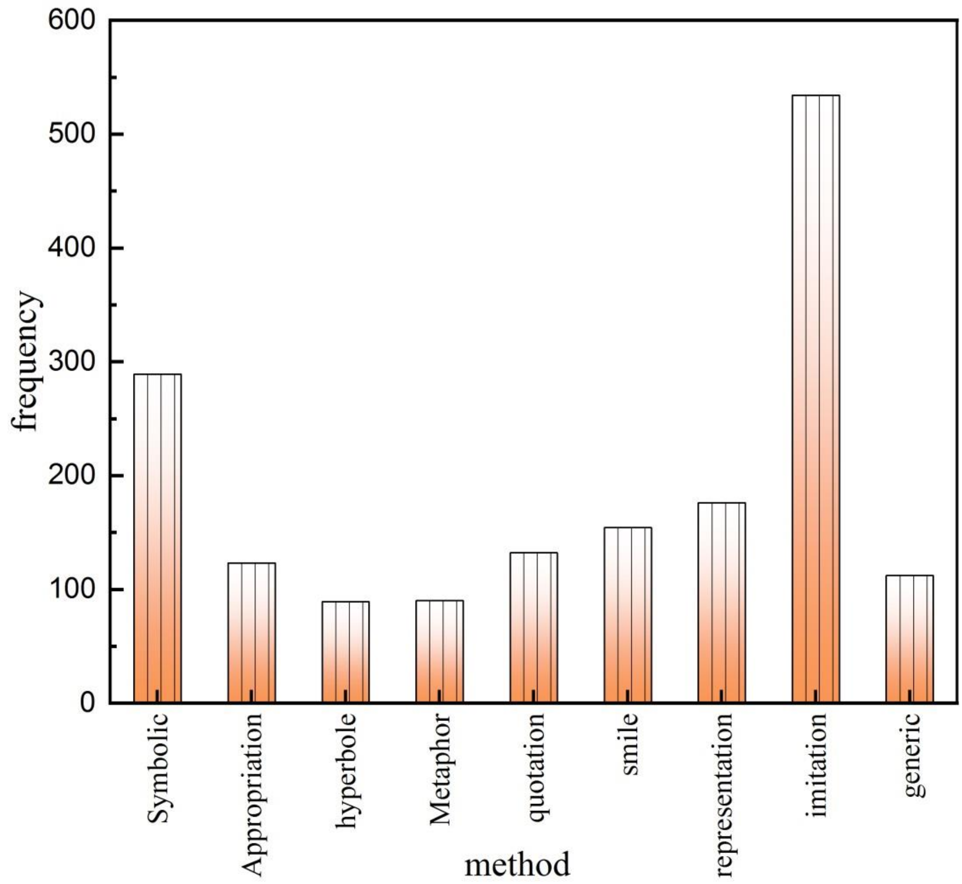



**Figure 6.** Image rhetorical statistics

### Visual segment rhetoric of Luoshan shadow based on improved VSLAM algorithm

A visual paragraph, equivalent to a paragraph in a verbal essay, a paragraph consists of sentences or groups of sentences, which are textual pauses in the expression of the content of the ideas of the essay due to transitions, emphases, intervals and other circumstances. Each paragraph will have a relatively complete explanation of a concept or thing and will assume a different content role. In a complete visual text, the conception and expression of the visual paragraph plays the same key role in coordinating the relationship of images and reflecting the hierarchy of discourse. Many illustration methods common to the illustrative genre are still reflected in visual texts. We can classify the rhetorical methods of visual segments into analogy, example, accumulation, comparison, repetition, pun, reference, comparison, etc., based on the different narrative forms embodied in the shadow texts, and there are many works that tend to narrate uninterruptedly according to the logic of the development of things, or narrate the content along the imagery or contextual vein one sentence at a time, exactly like poetry, and some texts also contain more than

Volume: 8, No: 1, pp. 434-452
ISSN: 2059-6588 (Print) | ISSN 2059-6596 (Online)

one visual segment rhetoric. As far as possible, one or two of the most common and typical rhetorical styles in each piece were selected. In the shadow works, there are 87 pieces of works that use anthropomorphic rhetoric, and there are about 100 pieces of works that use repetition, pun, reference, and exaggeration rhetoric. Examples account for the largest proportion, with about 220 pieces. Accumulation and comparison rhetoric have 100 pieces in shadow works. Figure 7 shows the rhetorical techniques of visual discourse.

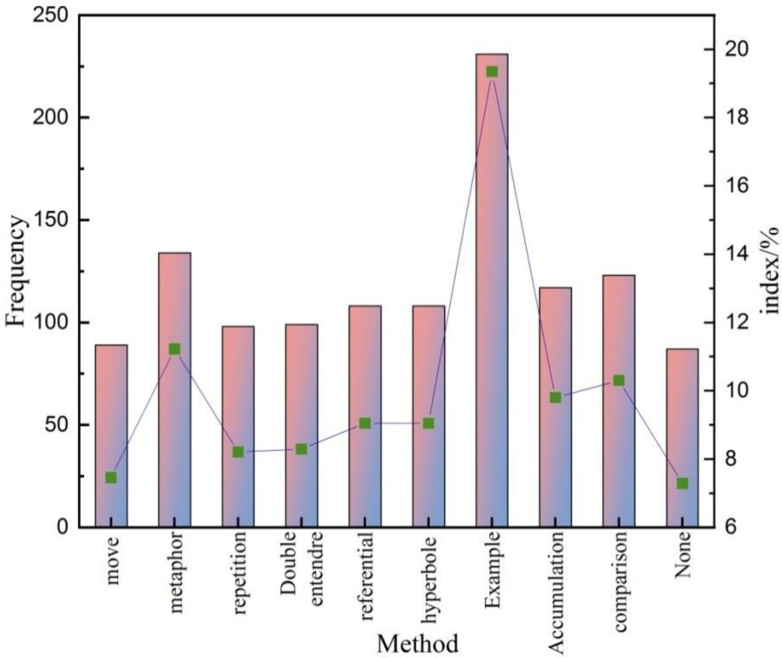



**Figure 7.** Visual discourse rhetorical devices

## Conclusion

In this paper, we analyze the symbolic representations and the rhetoric of images and visual segments used in the Luoshan shadow by VSLAM algorithm, and extract and analyze these representations and rhetoric, and compare the accuracy of representation extraction by VSLAM algorithm and improved VSLAM algorithm. The following conclusions are drawn:

The representation extraction time of the improved VSLAM algorithm is between 15ms-20ms, and that of the original VSLAM algorithm takes about 22ms. The feature values extracted by the improved algorithm are basically around 1000-2500, while the feature extraction values of the VSLAM algorithm are around 500-1500, and the improved VSLAM algorithm extracts more features than the original algorithm.

The matching accuracy of the original VSLAM algorithm features was between 40%-50%, and the improved algorithm matching accuracy was between 70%-80%, and the improved VSLAM algorithm accuracy was all improved by about 20%.

In terms of image rhetoric, the rhetorical technique of facsimile was used in 500 of the surveyed works. This was followed by the rhetorical device of symbolism, with a frequency of about 300. This was followed by anthropomorphism, representation, and quotation, with frequencies ranging from 120 to 150, respectively.

In terms of visual segmental rhetoric, 87 of the works use anthropomorphic rhetoric, while the number of works using repetition, pun, reference, and exaggeration rhetoric is generally around 100. Examples account for the largest number, with about 220 pieces. There are 100 pieces of accumulative and comparative rhetoric in shadow works.

**Funding**

This research was supported by Special Project Funding for Henan Province Philosophy and Social Science Planning (2022XWH237).

## References

A, I. (2016). Chinese Shadow : Figure Plays and Shadow Puppets. *Journal of the Faculty of Humanities University of Toyama, 19*, 34.

Binyon, H. (1966). *Puppetry Today; Designing and Making Marionettes, Hand Puppets, Rod Puppets, and Shadow Puppets*: New York: Watson-Guptil Publications.

Bryant, J., Heard, C., Watson, L., & MacDonald, M. (2002). *Making shadow puppets*: Kids Can Press Buffalo, NY.

Currell, D. (2015). *Shadow puppets and shadow play*: Crowood.

E, C. (2016). Making shadow puppets. *Practical Pre-School, 1*(10), 2

H, L. J. (2018). Study on the shape,color and craft characteristics of Luoshan shadow play. *Journal of Jiamusi Education Institute*.

H, Q. (2018). The Innovative Application of Xinyang Luoshan Shadow Play in Digital Animation. *Idea & Design*.

L, C. G. (2019). Contemporary art education value and the thinking of the traditional folk art. *Journal of Nanchang College of Education*.

M, Z. Q. (2015). On the Inheritance and the Development Historical Mechanism of Luoshan Shadow Play. *Journal of Xinyang Normal University (Philosophy and Social Sciences Edition)*.

Mendiola A , T. E. M. Without a Shadow of a Doubt: Shadow Puppets Lighting the Way to Literacy.

Özcan, O. (2002). Cultures, the traditional shadow play, and interactive media design. *Design Issues, 18*(3), 18-26.

Q, Z. (2018). On Natural Mode of the Inheritance and the Development of the Luoshan Shadow Play before it Becomes the Intangible Cultural Heritage. *Journal of Henan Institute of Education(Philosophy and Social Sciences)*.

Qin H W , D. A., University W A (2015). On the Fusion of Traditional Folk Art and Modern Design Art. *Journal of Zunyi Normal College*.

S, W. (2019). Shadow Shadow - Shadow Puppets and Shadow Puppet Plays [Book Review]. *Educating Young Children: Learning and Teaching in the Early Childhood Years, 15*(1).

Y, C. ( 2016). The Present Situation and Protection Measures of Shadow Play in South Henan. *Journal of Henan University of Science & Technology(Social Science)*.

Y, C. ( 2017). The Present Situation and Protection Measures of Shadow Play in South Henan. *Journal of Henan University of Science & Technology(Social Science)*.

Y, S. (2016). Cultures,the Traditional Shadow Play,and Interactive Media Design. *Journal of Nanping Teachers College*.

Yan C , A. S. O. U. C. N. (2016). Cultural Ecology of Traditional Folk Art and Cultural Characteristics and Values. *The Guide of Science & Education*.

Zhang J , L. F. W. (2017). Temple Fair:The Important Communicant Site of Folk Shadow Play. *Journal of West Anhui University*.