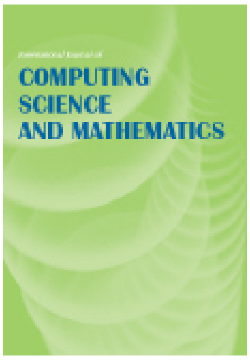


International Journal of Computing Science and Mathematics

ISSN online: 1752-5063 - ISSN print: 1752-5055
https://www.inderscience.com/ijcsm

Visual communication method for multi feature media images based on interactive modelling

Ting Zhang, Shi-shun Wang, Xiang-yu Wei

DOI: 10.1504/IJCSM.2024.10068569

Article History:

Received:	29 March 2024
Last revised:	12 August 2024
Accepted:	19 October 2024
Published online:	06 May 2025

Visual communication method for multi feature media images based on interactive modelling

Ting Zhang*

School of Design and Art,
Shanghai Dianji University,
Shanghai, 200240, China
Email: zhangting_sdju@163.com
*Corresponding author

Shi-shun Wang

School of Visual Communication Design,
Shandong University of Art and Design,
Jinan, 250300, China
Email: wangshishun0504@163.com

Xiang-yu Wei

School of Design and Art,
Shanghai Dianji University,
Shanghai, 200240, China
Email: XiangyuW@mls.sinanet.com

Abstract: In order to solve the problems of poor multi feature media image processing, low peak signal-to-noise ratio, and long visual communication processing time in traditional methods, a visual communication method for multi feature media images based on interactive modelling is proposed. hue-saturation-value (HSV) colour histogram, discrete Cosine transform (DCT) transform and generalised search tree (GIST) descriptor were used to extract colour, texture and scene content features of multi-feature media images, and the features were fused. The multi-feature media image is decomposed into semantically independent components, which are mapped to the surface of the 3D model, so as to realise the interactive modelling of the multi-feature media image. Combined with the modelling results, the image is enhanced and reconstructed to complete the visual transmission of the image. Experimental results show that the image detail features of the proposed method are significant, with high clarity. The average peak signal-to-noise ratio is 55.9 dB, and the processing time remains below 0.4 s.

Keywords: interactive modelling; multi feature media; visual communication; colour; texture; scene content; semantically independent components; enhanced.

Reference to this paper should be made as follows: Zhang, T., Wang, S-s. and Wei, X-y. (2025) ‘Visual communication method for multi feature media images based on interactive modelling’, *Int. J. Computing Science and Mathematics*, Vol. 21, No. 1, pp.14–31.

Biographical notes: Ting Zhang received her Bachelor's degree in Art Design from the School of Construction Engineering and Mechanics of Shanghai Jiaotong University; received her PhD in Architecture from School of Architecture and Urban Planning of Tongji University in 2013. Currently she is an Associate Professor in the School of Design and Art of Shanghai Dianji University. Her research interests include digital media art, interaction design, interface design and visual communication.

Shi-shun Wang received his Bachelor's degree in art design from Shandong University of Art and Design; received his PhD in Design Discipline from the School of Design and Innovation of Tongji University in 2021. Currently, he is a Lecturer in the School of Visual Communication Design of Shandong University of Art and Design. His research interests include visual communication, digital media art, interdisciplinary design, inclusive design system and urban design system.

Xiang-yu Wei received his Bachelor's degree in Electrical Engineering and Automation from the School of Electric Power, Civil Engineering and Architecture of Shanxi University in 2022. Currently, he is a master student in Mechanical College of Shanghai Danji University. His research interests include electrical engineering and automation, digital media art, interaction design, interactive modelling, intelligent algorithms, interface design and visual communication.

1 Introduction

With the continuous development of computer vision and artificial intelligence technology, the demand for visual communication research on multi feature media images is gradually increasing. In modern society, media communication and information rendering have become an indispensable part, and how to use multiple features to achieve image recognition and classification has become a research hotspot for scholars (Jiao, 2023). The research on visual communication of multi feature media images is an important field in computer vision and machine learning. Its applications are very extensive, including image classification, recognition, object detection, autonomous driving, and industrial intelligence (Zeng, 2022). The research on visual communication of multi feature media images refers to the use of multiple feature combinations to achieve image recognition and classification, providing technical support for media communication and information rendering (Huang, 2021). Specifically, research on visual communication of multi feature media images can improve visual impact, enhance information transmission efficiency, enhance interactivity, and adapt to the needs of the digital age. This method can be widely applied in fields such as image recognition and classification, virtual reality, etc., providing strong support for the development of computer vision and machine learning technology. Therefore, studying visual communication methods of multi feature media images is of great significance (Zhang and Peng, 2022).

In response to the above issues, Gao and Zhang (2021) proposed a new multi feature media image visual communication method based on laser shock. Using laser sensors to collect relevant images, using improved SIFI algorithm to determine image feature points, connecting feature points to obtain image contour lines, interpolating the obtained contour lines, determining the key and intermediate frames of multi feature media images, and combining laser shock method to construct multi feature media image feature sequences, and enhancing features to achieve visual communication processing of multi feature media images. However, in actual testing, it was found that this method has the problem of poor visual communication performance of multi feature media images. (Zhang, 2021) proposed a multi feature media image visual communication method based on spatiotemporal filtering. Utilising machine vision technology to collect multi feature media images, and implementing image noise suppression through guided filtering. Decompose the image to obtain an image sequence, and use partial differential equations to obtain the target points in the sequence, thereby achieving the goal of multi feature media image background suppression. The histogram equalisation method is used to balance and enhance the image, determine the information and characteristics of the target object, and optimise the image visual communication parameters with particle swarm optimisation algorithm to achieve the visual communication processing of multi feature media images. However, the peak signal-to-noise ratio of the multi feature media image processed by this method is low, and there is still a certain gap between it and the ideal target. Gardner et al. (2022) proposed a digital image visual communication method based on automated reconstruction of digital images. The CMOS protruding sensor is used to collect digital images, and the Bilateral filter method is used to filter the image noise component. The HOG feature is used to measure the gradient direction of the image, extract the texture features in the image, and enhance the image features, so as to achieve the automatic processing of digital images and visual communication processing. However, in practical applications, it has been found that this method has the problem of long visual communication time for digital images, and the actual application effect is not good.

In order to solve the problems of poor multi feature media image processing, low peak signal-to-noise ratio, and long visual communication processing time in traditional image visual communication methods, a visual communication method for multi feature media images based on interactive modelling is proposed.

2 Visual communication methods for multi feature media images

2.1 Image feature extraction and fusion

On the basis of previous research, the colour, texture, and scene content of multi feature media images were extracted (Man et al., 2022), and the features of three types of multi feature media images were normalised and fused to enhance the robustness of visual communication in multi feature media images. Image multi feature fusion technology can comprehensively utilise multiple features, fully extract data information, reduce computational complexity, effectively improve the robustness of multi feature media image visual communication processing, and make the image processing process more robust (Niu and Zhang, 2022).

1 *Colour feature extraction based on HSV colour histogram*

There are various methods for extracting colour features from multi feature media images. HSV colour histograms can distinguish colours in three dimensions: hue, saturation, and brightness, making it more accurate to distinguish different colours. Therefore, this paper uses this method to extract colours from multi feature media images (Tarsitano et al., 2022). Assuming there is a multi-feature media image point, mainly represented by (i, j), the colour feature extraction results of the image are as follows:

$$\overline{f}_I^{hsv} = H(i,j)\times 16 + S(i,j)\times 4 + V(i,j) \tag{1}$$

In the above formula, $H(i,j)$, $S(i,j)$, and $V(i,j)$ respectively represent the *H*, S, and *V* components at point (i,j) in the multi feature media image.

2 *Texture feature extraction based on DCT transform*

Discrete Cosine transform (DCT) has advantages in image texture feature extraction, such as good energy concentration, ease of calculation and implementation, good robustness, and ease of processing and analysis. The DCT transformation matrix is described using the following formula:

$$k_W = \begin{bmatrix} 1 & 1 & 1L & 1 & 1 \\ -N+1 & 1 & 1L & 1 & 1 \\ 0 & -N+2 & 1L & 1 & 1 \\ \vdots & \vdots & \vdots\,\vdots & \vdots & \vdots \\ 0 & 0 & 0L & 1 & 1 \\ 0 & 0 & 0L & -N+(N-1) & 1 \end{bmatrix} \tag{2}$$

In the above formula, N represents the maximum number of DCT waves.

Build a $3^2 \times 3^2$ DCT wave matrix based on the DCT transformation matrix constructed above, as shown in Figure 1.

Combining the DCT wave matrix obtained above, multiply the matrix parameters by the DCT transformation coefficient d to obtain the DCT wave. The calculation formula is as follows:

$$[F] = [k_W]\, d\, [k_W]^T \tag{3}$$

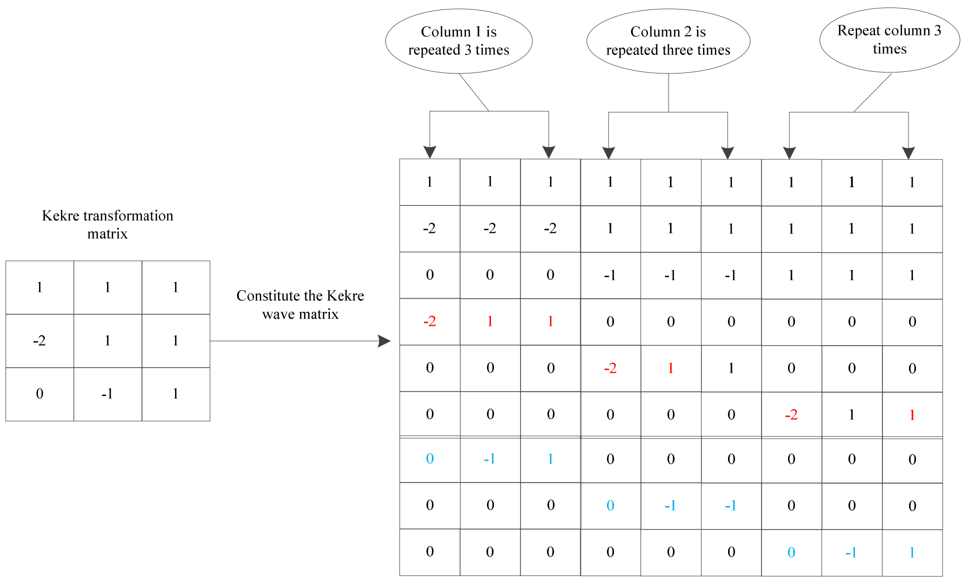
In the above formula, T represents matrix transpose.

Combining DCT waves to extract texture features of multi feature media images (Pei et al., 2021), the specific results are as follows:

$$\overline{f}_I^{dct} = \left\{0.4^* f_{DCT}, 0.6^* f_{DCT\_Wavelet}\right\} \tag{4}$$

In the above formula, f_{DCT} represents the DCT wave feature, and $f_{DCT\_Wavelet}$ represents the DCT feature vector.

Figure 1 DCT wave matrix (see online version for colours)



3 *Scene content feature extraction based on GIST descriptor*

In practice, GIST descriptors are mainly used to classify multi feature media image scenes, which is a global structural feature that can extract content features of multi feature media image scenes, laying a solid foundation for subsequent analysis. The process of extracting scene content features based on GIST descriptors is as follows:

1 Assuming there is a multi-feature media image $f(x, y)$ with a size of $h \times w$, divide $f(x, y)$ into $n_a \times n_a$ grids of equal size, represented by $h' \times w'$, where $h' = h / n_a, w' = w / n_a$.

2 If Gabor filter convolutional filtering is used to process each small block in a multi feature media image, the relationship of $n_c = m \times n$ holds, where m and n respectively represent the scale and number of directions of the filter (Park et al., 2022).

3 Calculate the feature values of each small image block and average them to obtain the GIST descriptor sub features of the multi feature media image block. The results are as follows:

$$\overline{G_{n_c}} = \frac{1}{h' + w'} \sum_{(x,y)} G_{n_c}(x, y) \tag{5}$$

In the above formula, the small feature values of the multi feature media image generated after filtering the n_c th channel of $\sum_{(x,y)} G_{n_c}(x, y)$.

4 Cascade the n_c average feature values generated from each small block in step (3) to obtain the scene content feature $\overline{f_I^{gist}}$ of the entire image, with a dimension of $n_a \times n_a \times n_c$.

Before fusing the three features, it is necessary to normalise the colour, texture, and scene content features. The specific normalisation formula is as follows:

$$\overline{f}_I^c(t) = \frac{f_I^c(t)}{\sum_t f_I^c(t)}, c \in \{hsv, dct, gist\} \tag{6}$$

In the above formula, $f_I^c(t)$ represents the t th component of multi feature media image feature c.

The fusion results of multi feature media image features are as follows:

$$F_I = \{s * \overline{f}_I^{hsv}, z * \overline{f}_I^{dct}, l * \overline{f}_I^{gist}\} \tag{7}$$

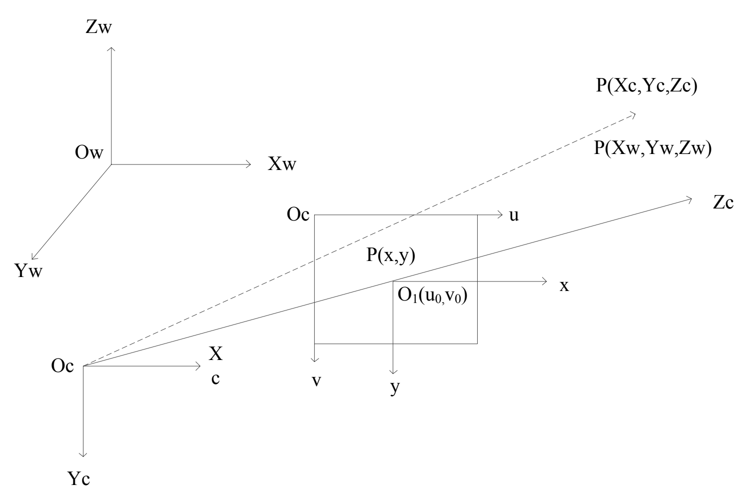
In the above formula, s represents colour feature weight, z represents texture feature weight, and $\overline{f}_I^{gist}$ represents scene content feature weight.

2.2 Interactive modelling of multi feature media images

Multi feature media image interactive modelling technology (Streeb et al., 2021) has improved modelling efficiency and experience, lowered modelling thresholds, and has been applied in multiple fields, mainly detecting image feature points based on feature fusion results and completing initial camera calibration. Decompose multi feature media images into semantically independent components, and map them to the surface of the 3D model based on the segmented component contours to achieve interactive modelling of multi feature media images (Wen et al., 2022).

Combining the feature extraction and fusion results of multi feature media images, determine image feature points, and use any feature point as the origin to establish a reference coordinate system of $O_W X_W Y_W Z_W$. The camera coordinate system is shown in Figure 2.

Figure 2 Camera coordinate system



Assuming the camera matrix is represented by $M = K[R|T], K, R, T \in R^3$, where R and T represent different external parameter matrices and K represent internal parameter matrices, the camera parameter calibration process is as follows:

1 *Internal parameters*

Assuming that the distortion parameter of the camera used in this paper is represented by s, and ideally this value needs to trend endlessly towards 0, the internal parameter calibration function can be represented by $f_x = f_y = f$ to obtain the internal parameter calibration result (Dehghan et al., 2021), which can be represented by (u_0, v_0).

2 *External parameters*

Assuming that ϕ, θ, φ represents the three Euler angles rotating around $X_W Y_W Z_W$ and t_x, t_y, t_z represents the three different matrix parameters of T, R can be represented by the following formula:

$$R = \begin{bmatrix} \cos\phi\cos\theta & \cos\phi\sin\theta\sin\varphi - \sin\phi\cos\varphi & \cos\phi\sin\theta\cos\varphi - \sin\phi\cos\varphi \\ \sin\phi\cos\theta & \sin\phi\sin\theta\sin\varphi + \cos\phi\cos\varphi & \sin\phi\sin\theta\cos\varphi + \cos\phi\sin\varphi \\ -\sin\theta & \cos\theta\sin\varphi & \cos\theta\sin\varphi \end{bmatrix} \tag{8}$$

Based on the camera model in Figure 2, the optical axis of the camera is perpendicular to the image plane, resulting in vanishing points v_1, v_2 in two orthogonal directions. These two points are represented by $(x_1, y_1, f), (x_2, y_2, f)$ in the coordinate system. According to the camera calibration method based on vanishing points, the focal length of the camera is calculated, and the results are as follows:

$$f = \sqrt{\left|(x_1 - u_0)(x_2 - u_0) + (y_1 - v_0)(y_2 - v_0)\right|} \tag{9}$$

Assuming that the coordinates of image feature points in the world coordinate system are mainly represented by (u, v), the external parameter matrix T is represented by the following formula:

$$\lambda \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = K[R, T] \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} = KT \tag{10}$$

In the above formula, (u, v) represents the pixel coordinates of feature points in multi feature media images, and λ represents the discrete coefficients.

After calibrating camera parameters, it is necessary to combine relevant image segmentation algorithms to segment multi feature media images, in order to obtain multiple semantic components with independent properties. Based on the obtained results, combined with relevant modelling primitives, three-dimensional fitting is performed on each component to achieve interactive modelling of multi feature media images. The Grab Cut algorithm is an intelligent algorithm used for image classification, which separates foreground and background in an image based on boundary separation between foreground and background. This algorithm achieves interactive image segmentation by manually or semi-automatically selecting regions of interest, thereby improving the accuracy of segmentation. Minimise the Gibbs energy function $E(\alpha,k,\theta,z)$ of the image through iterative optimisation:

$$E(\alpha,k,\theta,z)=U(\alpha,k,\theta,z)+V(\alpha,z) \tag{11}$$

In the above formula, $U(\alpha,k,\theta,z)$ represents the penalty function, and $V(\alpha,z)$ represents the edge feature fitting function of multi feature media images.

Multi feature media image target object modelling primitive fitting, the specific objective function is represented by the following formula:

$$\min E=\sum_{i=1}^{m} d_2\left(c_i,\text{outline}(p,M)\right)^2 \text{ subject to}:\phi(p)=0 \tag{12}$$

In the above formula, d_2 represents the distance between domain pixels, $\phi(p)$ represents the internal constraint parameters of the primitive, and c_i represents the feature points on the contour line.

The geometric semantic relationship between different modelling primitives is taken as a constraint, and global optimisation is carried out for the primitive fitting results of multi feature media image target objects. The objective function is expressed by the following formula.

$$\min E=\sum_{i=1}^{n} w_i\left(\sum_{j=1}^{m} d_2\left(c_j,\text{outline}(p_i,M)\right)^2\right)\text{subject to}:\phi(p_i)=0 \tag{13}$$

In the above formula, w_i represents the weight of the *i*th primitive component, and c_j represents the number of feature points of the primitive component.

Assuming that each modelling primitive *p* is defined by a set of parameters x_p, the objective function $f_p(x_p)$ is minimised under the constraint of the internal structural relationship $\phi_p(x_p)=0$ of the primitive to achieve the fitting of primitive p with the target object component. The constraint relationship between primitives is represented by a set of geometric semantic relationships G. For a geometric semantic relationship $g\in G$, define x_g as the parameter in the relationship, and define the constraint of its geometric semantic relationship as $\psi_g(x_g)=0$. Note that for the geometric semantic relationship g between a set of primitives, parameter x_g should be a subset of $\bigcup_p x_p$. If $P=\{p_1,\dots,p_n\}$ is the currently modelled *n* primitives, its corresponding parameter vector is represented as $x=\{x_{p_1},\dots,x_{p_n}\}$. After primitive *p* and the multi feature media image contour are fitted to generate the initial model, the geometric semantic relationship between primitives is detected according to x_p^* and added to constraint set G, and the generated new primitive *p* is added to set *P* to global optimisation the objective functions of all primitives:

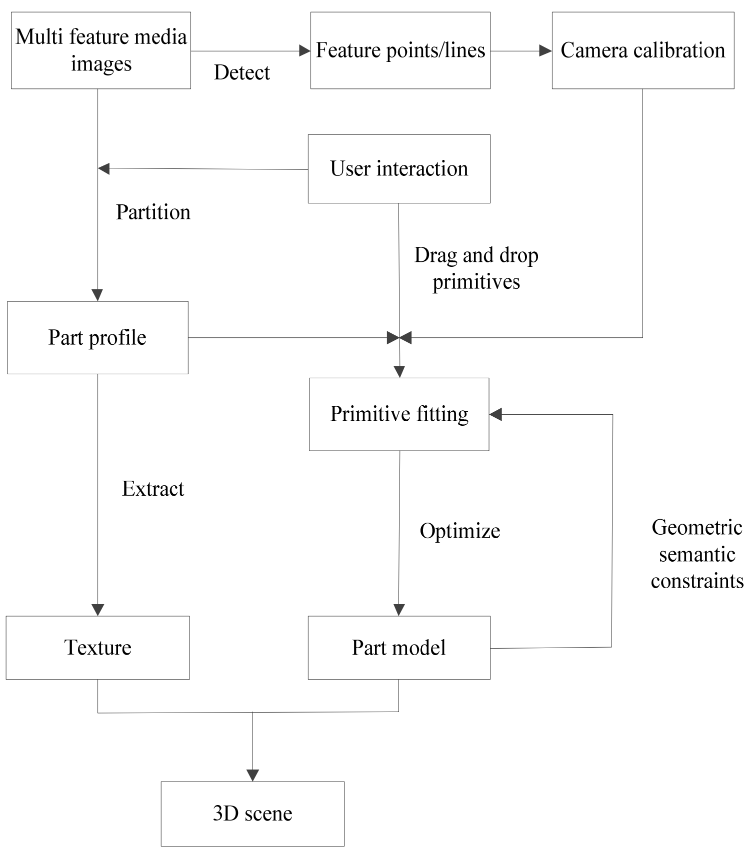
$$\begin{cases} \min : \sum_{p \in P} f_p\left(x_p\right) \\ s.t.: \phi_p\left(x_p\right) = 0 \forall_p \in P \\ \Psi_g\left(x_g\right) = 0 \forall_g \in G \end{cases} \tag{14}$$

Let x_{prev}^* be the optimal parameter for the previously modelled primitive generated by the new primitive p, and define the initial point as $x_{start} = \left(x_{prev}^*, x_p^*\right)$, assuming that there is no geometric semantic constraint between the new primitive p and the primitive at the initial time. The augmented Lagrangian method solves the problem by transforming the constrained optimisation problem into a series of unconstrained problems. Given constants μ and $\lambda = \{\lambda_1, ..., \lambda_m\}$, construct a multi feature media image interactive model, as follows:

$$Y = f\left(x\right) + \sum_{i=1}^{m} \lambda_i \cdot c_i\left(x\right) + \frac{\mu}{2} \sum_{i=1}^{m} \left|c_i\left(x\right)\right|^2 \tag{15}$$

The interactive modelling process of multi feature media images is shown in Figure 3.

Figure 3 Interactive modelling process of multi feature media images



2.3 Visual communication based on image reconstruction

By combining interactive modelling results and utilising guided filtering and Laplacian pyramid for image enhancement and reconstruction, a multi feature media image visual communication design is completed. Its innovation lies in the fact that guided filtering is a local smoothing filtering method based on image gradients. Its basic idea is to connect gradient information and image smoothing, which can effectively remove noise from images Preserve edge information and improve the visual quality of multi feature media images. The Laplace pyramid can help extract noise information from images, and noise becomes more pronounced at high frequency levels. By filtering the high-frequency layers, the impact of noise on the image can be effectively suppressed. Combining guided filtering and Laplacian pyramid can ensure the quality of multi feature media image reconstruction. Therefore, this method has multiple characteristics such as good multi feature media image reconstruction effect and fast speed, which can improve the visual communication effect of multi feature media images, better recover damaged image areas, eliminate noise, etc., improve image contrast, etc. This is of great significance for fields such as digital image processing, image restoration, and enhancement.

Using guided filtering method to perform linear transformation on multi feature media images, one guided image I_i, one filtered input image p_i, and one output image q_i are obtained. The formula for calculating the filtering value of pixel i in multi feature media images is as follows:

$$q_i = a_k I_i + b_k, \forall_i \in \omega_k \tag{16}$$

Among them, (a_k, b_k) represents the linear coefficient and bias, respectively, and is a constant in window ω_k with a radius of r. Guided filtering has good edge preservation performance, as $\nabla q_i = \mathrm{a}\nabla I_i$, only when image I_i has edge details, can q_i use edges.

To determine the linear coefficient a_k and bias b_k, constraints from filter input image p_i are required. The value of output image q_i can be modelled as input image p_i minus noise component *n*. The specific expression is as follows:

$$q_i = p_i - n_i \tag{17}$$

The output multi feature media image is solved by minimising the loss function in window ω_k, where the expression of the loss function is as follows:

$$E(a_k, b_k) = \sum_{\mathrm{i}\in\omega_\mathrm{k}} \left((a_k I_i + b_k - p_i)^2 + \varepsilon a_\mathrm{k}^2 \right) \tag{18}$$

where ε represents the coefficient of the regularisation penalty term of the linear coefficient a_k, and ω_k represents the total number of pixels. The solution of this linear regression model is as follows.

$$a_k = \frac{\frac{1}{|\omega|}\sum_{i\in\omega_k} I_i p_i - \mu_k \overline{p}_k}{\sigma_k^2 + \varepsilon}, b = \overline{p}_k - a_k \mu_k \tag{19}$$

In the above formula, μ_k represents the loss coefficient and $\overline{p}_k$ represents the pixel mean.

Although image filtering and enhancement processing can be achieved through the above process, some details in multi feature media images may be lost during this process. The Laplace pyramid can supplement some missing details in the guided image and enhance the edges of the denoised image. Assuming that the noise gradient in greyscale image I_r can be removed through a Wiener filter. The formula is as follows:

$$\hat{f}(x,y) = g(x,y) - \frac{\sigma_g^2}{\sigma_L^2}\left[g(x,y) - \hat{\mu}_L\right] \tag{20}$$

In the above formula, $g(x,y)$ represents the pixel value of greyscale image I_r, σ_g^2 represents the mean of all pixels in the image, σ_L^2 represents the variance of different centre pixels (x,y), and $\hat{\mu}_L$ represents different centre pixels (x,y). $\hat{\mu}_L$ and σ_L^2 are calculated using the following formula:

$$\hat{\mu}_L = \frac{1}{MN}\sum_{(x,y)\in\Omega} g(x,y) \tag{21}$$

$$\sigma_L^2 = \frac{1}{MN}\sum_{(x,y)\in\Omega} g^2(x,y) - \hat{\mu}_L{}^2 \tag{22}$$

In the above formula, M and N are the height and width of the window centred on (x,y).

After preprocessing with the Wiener filter, the three channel images processed by the filter are synthesised into RGB image I_{wiener}. At this point, the noise level of the RGB image processed by the filter has been reduced to a certain extent. In order to protect the inherent texture and remove the incorrectly identified texture, the TV regularisation model is used to smooth the I_{wiener} multi feature media image. Define I_{wiener} as the input image and P as the output image. The solution method for *P* is as follows.

$$\min_s \sum_k \left(P_k - I_{wiener-k}\right)^2 + \lambda^* C(P) \tag{23}$$

In the above formula, $\left(P_k - I_{wiener-k}\right)^2$ represents the fidelity term, *k* represents the order of pixels in P and *I*, and $C(P)$ represents the regularisation term. The calculation method is as follows.

$$C(P) = \left\{k\left|\partial_x P_x\right| + \left|\partial_y P_x\right| \neq 0\right\} \tag{24}$$

By solving and smoothing P, the texture caused by noise is removed, but the quality of the multi feature media image still cannot reach the ideal state. Therefore, by enhancing the details of the multi feature media image through Gaussian differential filtering, the visual communication effect of the multi feature media image is improved. Gaussian kernels of different scales are applied to global image I, resulting in three different blurred images:

$$\begin{cases} B_1 = G_1 * I \\ B_2 = G_2 * I \\ B_3 = G_3 * I \end{cases} \tag{25}$$

In the above formula, G_1, G_2, and G_3 represent different standard Gaussian kernels. Through the above process, fine details D_1 and intermediate details $D_2 D_3$ of multi feature media images are extracted. The specific calculation formula is as follows:

$$\begin{cases} D_1 = I - B_1 \\ D_2 = B_1 - B_2 \\ D_3 = B_2 - B_3 \end{cases} \tag{26}$$

Merge the three to obtain the overall detailed image, and the results are as follows:

$$D = \left[1 - W_1 \cdot \operatorname{sgn}\left(D_1\right)\right] \cdot D_1 + W_s \cdot D_2 + W_t \cdot D_3 \tag{27}$$

In the above formula, W_1 represents the Laplacian contrast weight, W_s represents the significance weight, and W_t represents the saturation weight. In the visual communication process of multi feature media images, it is necessary to merge three weight maps into a separate weight map. To achieve this goal, it is necessary to sum the three weight maps to obtain the sum weight map. The specific calculation formula is as follows:

$$W_k = W_1 + W_s + W_t \tag{28}$$

The goal of normalising the weight map of a multi feature media image is achieved by dividing the weight of each pixel in the image by the sum of the weights of the same pixels in all images. The result is represented by the following formula:

$$\bar{W}_k = \left(W_k + \delta\right) / \left[\sum_{k=1}^{k} W_k + k \cdot \delta\right] \tag{29}$$

In the above formula, δ represents the regularisation term, and the visual transmission function of multi feature media images is designed by using the Laplacian pyramid to system the quality of multi feature media images. The definition of the Laplace pyramid is as follows:

$$\begin{aligned} I(x) &= I(x) - G_1\{I(x)\} + G_1\{I(x)\} \to L_1\{I(x)\} + G_1\{I(x)\} \\ &= L_1\{I(x)\} + G_1\{I(x)\} - G_2\{I(x)\} + G_2\{I(x)\} \\ &= L_1\{I(x)\} + L_2\{I(x)\} + G_2\{I(x)\} \\ &= \cdots \\ &= \sum_{l=1}^{N} L_l\{I(x)\} \end{aligned} \tag{30}$$

In the above formula, L_l and G_l represent the l th layer of the Laplace pyramid and the Gaussian pyramid, respectively.

The visual communication function of multi feature media images is as follows:

$$R_l(x) = \sum_{k_x=1}^{k_x} G_l\left[\bar{W}_{k_x}(x)\right] \cdot L_l\left[I_{k_x}(x)\right] \tag{31}$$

In the above formula, l represents the number of pyramid layers, and k_x represents a sequence of multi feature media images.

3 Experimental scheme and results

3.1 Experimental scheme

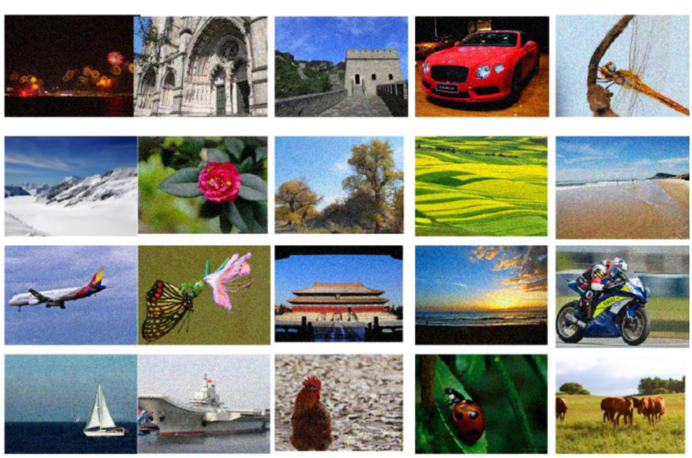
The research objective was to verify the effectiveness of the visual communication method for multi feature media images based on interactive modelling proposed in this paper. Experimental tests were conducted, and all experimental results were obtained in accordance with the experimental plan. Therefore, the experimental plan was designed with emphasis, as follows:

1 The experiment was conducted in a Windows 10 system environment, with Intel Core i5-750 processor, 12GB DDR2 667MHz memory, 256GB hard drive, Visual Basic programming language, and Matlab simulation software.

2 *Experimental data*: At this time, all experimental data comes from the GHIM-10K database, which contains a total of 10,000 images, including fireworks, architecture, the Great Wall, sports cars, dragonflies, snowcapped mountains, flowers, trees, grasslands, beaches, airplanes, butterflies, the Forbidden City, sunrise, motorcycles, sailboats, ships, hens, beetles, horses, and other categories. The images in the GHIM-10K database are evenly divided into a test set and an experimental set. The data from the test set is input into the simulation software for testing, and the simulation software parameters are adjusted through multiple tests until they are optimal. The adjusted simulation software is then subjected to subsequent testing to ensure the authenticity and reliability of the experimental results.

Some experimental sample images are shown in Figure 4.

Figure 4 Partial experimental sample images (see online version for colours)



3 *Evaluation indicators*

The evaluation indicators include the visual communication effect of multi feature media images, the peak signal-to-noise ratio of multi feature media images, and the visual communication processing time of multi feature media images as experimental indicators. The visual communication effect of multi feature media images refers to the level of image quality after image processing. The higher the image quality, the better the visual communication processing effect of the image; The peak signal-to-noise ratio of multi feature media images is an indicator used to evaluate image quality. It is the logarithmic signal-to-noise ratio of the maximum allowable value of the evaluation image to the ratio of noise. The higher the value, the better the image processing effect; The shorter the visual communication processing time of multi feature media images, the higher the visual communication processing efficiency.

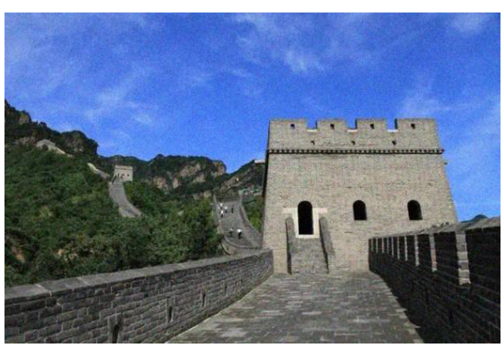
3.2 Experimental result

3.2.1 Visual communication effect of images

Select one image in the experiment as a representative, and use the method of this paper, Method of Gao et al. (2021), and Method of Zhang et al. (2021) to visually convey the image. The visual communication effect of multi feature media images is shown in Figure 5.

Figure 5 Visual communication effects of multi feature media images: (a) method of Gao et al. (2021); (b) method of Zhang et al. (2021) and (c) method of this paper (see online version for colours)

(a)



(b)

Figure 5 Visual communication effects of multi feature media images: (a) method of Gao et al. (2021); (b) method of Zhang et al. (2021) and (c) method of this paper (see online version for colours) (continued)

(c)

Analysing the results in Figure 5, it can be seen that Method of Gao et al. (2021) still has a relatively blurry image after visual communication processing for multi feature media images, making it difficult to determine image features. The image quality and clarity are always at a low level, indicating that the visual communication processing effect of this method is not good. However, Method of Zhang et al. (2021) still contains a lot of noise components in multi feature media images after visual communication processing, resulting in poor image clarity and processing effectiveness. After processing the image using this method, the image details are more prominent and the image clarity is higher, indicating that the visual communication processing effect of this method is higher. The reason is that this method combines interactive modelling results with guided filtering and Laplacian pyramid for image enhancement and reconstruction processing, completing the visual communication design of multi feature media images. Therefore, the practical application effect of this method is relatively good.

3.2.2 Peak signal-to-noise ratio

Using other images for multi feature media image visual communication processing experiments, the comparison results of the peak signal-to-noise ratio of the three methods are shown in Table 1.

From Table 1, it can be seen that the maximum peak signal-to-noise ratio of multi feature media images in Method of Gao et al. (2021) is 47.5 dB, and the maximum value in Method of Zhang et al. (2021) is 38.6 dB. The maximum value in this method is 57.3 dB, which is 9.8 dB and 18.7 dB higher than Method of Gao et al. (2021) and Method of Zhang et al. (2021), respectively; The minimum peak signal-to-noise ratio of multi feature media images in Method of Gao et al. (2021) is 35.6 dB, while the minimum value in Method of Zhang et al. (2021) is 31.7 dB. The minimum value in this method is 53.6 dB, which is 18 dB and 22.9 dB higher than Method of Gao et al. (2021) and Method of Zhang et al. (2021), respectively; The average peak signal-to-noise ratio of multi feature media images in Method of Gao et al. (2021) is 41.1 dB, while the average value in Method of Zhang et al. (2021) is 35.2 dB. The average value in this method is 55.9 dB, which is 14.8 dB and 20.7 dB higher than Method of Gao et al. (2021) and Method of Zhang et al. (2021), respectively. From all aspects, the peak signal-to-noise ratio of this method is the highest, indicating that its image quality is higher and its visual communication effect is better. The reason is that this method combines guided filtering and Laplacian pyramid for multi feature media image enhancement and reconstruction processing, which can fully utilise the structural information and frequency domain features of the image in image enhancement and reconstruction, while suppressing noise. This can maintain the details, textures, and other features of the image during the process of image enhancement and reconstruction, while reducing the impact of noise.

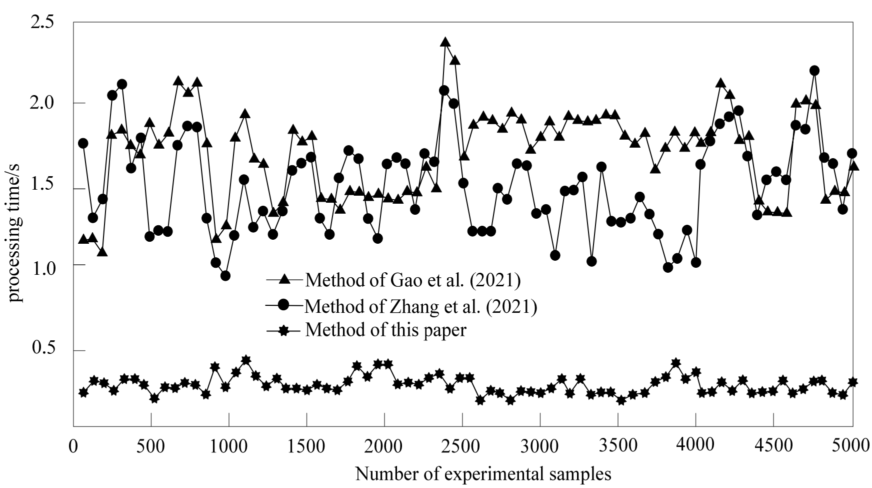
Table 1 Comparison of peak signal-to-noise ratio of images (Unit: dB)

Number of experimental samples	*Method of Gao et al. (2021)*	*Method of Zhang et al. (2021)*	*Method of this paper*
500	42.3	32.4	56.3
1000	35.6	31.7	55.8
1500	35.9	38.6	54.7
2000	36.8	35.5	56.9
2500	42.7	34.7	53.6
3000	41.2	36.9	55.8
3500	43.2	38.6	54.2
4000	47.5	35.2	56.8
4500	44.8	33.5	57.2
5000	41.6	34.9	57.3
Mean value	41.1	35.2	55.9

3.2.3 Image visual communication processing time

The comparison results of visual communication processing time for multi feature media images using three methods are shown in Figure 6.

Figure 6 Visual communication processing time



Analysing the results in Figure 6, it can be seen that the visual communication processing time of multi feature media images in Method of Gao et al. (2021) varies between 1.1 s and 2.3 s, while the visual communication processing time of multi feature media images in Method of Zhang et al. (2021) varies between 1.8 s and 2.1 s. However, the visual communication processing time of multi feature media images in this method always varies below 0.4 s, indicating that the visual communication processing time of multi feature media images in this method is shorter, higher efficiency and good practical application results. The reason is that this method uses the extraction and fusion of multi feature media image features, decomposing the multi feature media image into semantically independent components, and mapping them to the surface of the 3D model based on the segmented component contours, thereby achieving interactive modelling of multi feature media images. By combining the results of interactive modelling for image enhancement and reconstruction, a multi feature media image visual communication method is designed, which has high image processing efficiency.

4 Conclusions

As one of the important research fields in computer vision and machine learning, this research can achieve image recognition and classification through the combination of multiple features, and can be widely applied in multiple fields. It can solve problems such as image classification and recognition that people encounter in daily life, providing more efficient and accurate services for people, so this paper proposes a visual communication method for multi feature media images based on interactive modelling. The experimental results show that after processing the image using the method proposed in this paper, the image details are more prominent, the image clarity is higher, the average peak signal-to-noise ratio of the image is 55.9 dB, and the visual communication processing time of the image always changes below 0.4 s. It has the characteristics of good visual communication processing effect and high efficiency, and can be widely applied in the field of image processing. In the future, research on visual communication of multi feature media images will face greater challenges, such as how to improve the accuracy of image processing algorithms, how to solve the challenges of big data processing, and how to make algorithms more efficient. Therefore, future research on visual communication of multi feature media images needs to continuously improve and innovate by combining emerging technologies such as big data analysis, machine learning, and cloud computing.

References

Dehghan, M.J., Ardestani, V.E. and Dehghani, A. (2021) 'Moho topography estimation using interactive forward modeling of gravity data', *Journal of the Earth and Space Physics*, Vol. 32, No. 1, pp.122–138.

Gao, Y. and Zhang, J. (2021) 'The visual communication design of dim small target image based on time-space filtering', *Computer Simulation*, Vol. 38, No. 5, pp.168–171, 208.

Gardner, W., Winkler, D.A., Cutts, S.M., and Ardestani, V. (2022) 'Two-dimensional and three-dimensional time-of-flight secondary ion mass spectrometry image feature extraction using a spatially aware convolutional autoencoder', *Analytical Chemistry*, Vol. 94, No. 22, pp.7804–7813.

Huang, W. (2021) 'Design of 3D image visual communication optimization system based on digital image reconstruction', *Modern Electronics Technique*, Vol. 44, No. 4, pp.164–167.

Jiao, X.J. (2023) 'Visual communication method of ship image in fog based on spatio-temporal filtering', *Ship Science and Technology*, Vol. 45, No. 3, pp.173–176.

Man, H.S., Kim, C., Jin, G.H., Hu, L. and Qin, K. (2022) 'Image feature extraction and fractality evaluation based on two-dimensional continuous wavelet transform: application to digital elevation model data', *Fractals*, Vol. 30, No. 6, pp.1–12.

Niu, C. and Zhang, M. (2022) 'Using image feature extraction to identification of ancient ceramics based on partial differential equation', *Advances in Mathematical Physics*, Vol. 15, No. 6, pp.1–12.

Park, J., Kim, Y., Na, K. and Jin. G.H. (2022) 'An image-based feature extraction method for fault diagnosis of variable-speed rotating machinery', *Mechanical Systems and Signal Processing*, Vol. 167, No. 1, pp.108524–108536.

Pei, L., Yu, T., Yuan, H., Luo, G. and Wen, X.X. (2021) 'Aggregate weight prediction based on two-dimensional image feature extraction', *IOP Conference Series Earth and Environmental Science*, Vol. 668, No. 1, pp.012069–012081.

Streeb, D., Metz, Y., Schlegel, U., and Bruderer, C. (2021) 'Task-based visual interactive modeling: decision trees and rule-based classifiers', *IEEE Transactions on Visualization and Computer Graphics*, Vol. 28, No. 9, pp.3307–3323.

Tarsitano, F., Bruderer, C., Schawinski, K., Dehghani, A. and Winkler, D.A. (2022) 'Image feature extraction and galaxy classification: a novel and efficient approach with automated machine learning', *Monthly Notices of the Royal Astronomical Society*, Vol. 511, No. 3, pp.3330–3338.

Wen, Y.P., Hou, J.J., Tan, Z., Zhang, K. and Yuan, Y. (2022) 'Scientific workflow activity recommendation for interactive modeling', *Computer Integrated Manufacturing Systems*, Vol. 28, No. 10, pp.3115–3121.

Zeng, L.J. (2022) 'Research on virtual reconstruction of 3D image with complex illumination based on visual communication technology', *Laser Journal*, Vol. 43, No. 4, pp.93–97.

Zhang, W. (2021) 'Design of 3D image visual communication system based on automatic digital image reconstruction', *Manufacturing Automation*, Vol. 43, No. 6, pp.103–106, 141.

Zhang, Y.T. and Peng, T.Y. (2022) 'Design of multi-frame dynamic image visual transmission method based on laser shock', *Laser Journal*, Vol. 43, No. 3, pp.123–127.