International Journal for Housing Science and Its Applications
Publish August 10, 2025. Volume 46, Issue 4 Pages 4876-4885

https://doi.org/10.70517/ijhsa464331

Optimization Algorithm Solution and Accuracy Improvement for Clothing Size Matching Problems

Mengmeng Hou[1,*]
[1] Department of Fashion and Apparel Design, Zhengzhou Academy of Fine Arts, Zhengzhou, Henan, 451450, China
Corresponding authors: (e-mail: 13703848103@163.com).

Abstract This paper proposes a high-precision solution framework for garment size matching based on an optimization algorithm, which integrates binocular vision and feature extraction techniques to significantly improve the measurement accuracy. Firstly, the camera calibration is realized by using Zhang's calibration method combined with Brown's distortion model. And the innovative fusion of SIFT and Forstner algorithms optimizes the feature points, SIFT prescreening reduces the computational amount, and Forstner finely locates the corner points by error ellipse circularity thresholding. Further, the SGBM dense stereo matching algorithm is used to generate high-precision parallax maps by overcoming the “tail-dragging effect” through Sobel edge enhancement, multi-directional dynamic planning and post-processing. Experimental validation shows that the average error of the virtual grid calibration method at 10 positions is only 0.13mm, and the maximum error is 0.32mm, which meets the requirement of measurement error <1mm. In the DeepFashion dataset 13 types of clothing test, the average mAP reaches 77.37%, which is 4 percentage points higher than MKMnet, and the long-sleeved top has the best accuracy of 92.49%. The average matching accuracy under rotational interference is 88.46%, 22.69 percentage points higher than MKMnet, and the time consumed is 324ms, with an efficiency improvement of 7.2%.

Index Terms garment size matching, binocular vision, SGBM stereo matching, feature point detection, camera calibration

I. Introduction

Nowadays, the apparel industry, as a rigid industry to satisfy people's daily life, gradually transforms to the digital trend. Many Internet-based apparel platforms have appeared, and present a small batch, multi-species, personalized apparel customization mode [1]. This mode has become the development trend and new normal in the clothing industry, which is a kind of clothing industrialization mode that focuses on customers and realizes the whole process of customers from clothing design, to digital production, and then to service with the help of network platform [2]-[4]. In the online platform, consumers do not have an in-depth grasp of anthropometric measurements and their own body shape, which leads to a misjudgment of clothing size and an inappropriate understanding of whether the clothing layout is suitable for them [5]-[7]. This leads to the customization efficiency of the clothing industry is often relatively low, often appear the phenomenon of return and exchange, and consume a lot of human and material resources [8].

Clothing customization has the characteristics of diversification, personalization, and high-grade, which requires clothing companies need to be based on different customer needs, anthropometric data and clothing specifications for the “local” digital matching, so as to optimize the allocation of production resources, in order to improve the digital production technology of clothing customization [9]-[12]. At present, most of the domestic and foreign customized clothing enterprises are in the stage of scientific management, integrating brand design and production and processing, and have better execution in the informatization of customized clothing data processing and management concepts [13]-[15]. Therefore, through certain optimization algorithms to scientifically, accurately and adequately assess the degree of matching between garment sizes and user data, optimize the production process of garment customization enterprises, and maximize the production efficiency are the increasingly prominent production challenges of current garment customization enterprises.

This paper details the numerical solution scheme for the key techniques in the garment size matching problem based on the optimization algorithm, aiming to improve the overall measurement accuracy. First, the lens aberration is eliminated and an accurate imaging geometry model is established through high-precision camera calibration. The classical Zhang's calibration method combined with Brown's aberration model is used to solve the internal, external and aberration parameters of the camera using a planar tessellated lattice. Secondly, the focus is on solving

the efficiency and accuracy problems of garment feature point detection. The scale-space extreme point detection of SIFT algorithm is used as a fast initial screening mechanism, which effectively reduces the number of points that need to be processed by Forstner algorithm. On the initial point set, the autocorrelation matrix is calculated by using the gray scale variation in the neighborhood of the pixel points, and the corner points are accurately located by the roundness threshold and interest value of the error ellipse. This SIFT pre-screening + Forstner fine localization strategy significantly improves the speed and reliability of feature point extraction. Finally, the SGBM dense stereo matching algorithm is deeply applied to reconstruct the 3D information of the garment surface from the corrected binocular image.SGBM enhances the image edges by preprocessing with the Sobel operator, combines the gradient information and the original pixel information to compute the matching cost, and employs a multi-directional dynamic planning strategy to minimize the global energy function, which effectively overcomes the traditional dynamic planning's “trailing effect”.

II. Realization of key technology for garment size matching based on binocular vision

II. A. Clothing image preprocessing - camera calibration

In the process of image measurement, in order to determine the interrelationship between the three-dimensional spatial position of a point on the surface of an object in a real scene and its corresponding point in the image, it is necessary to establish a geometrical model of the camera imaging. These geometric model parameters are camera parameters, which need to be obtained through experiments and calculations, and the process of solving the camera parameters is the so-called camera calibration.

The lens of a camera is equivalent to a convex lens, and the light emitted from an object is focused onto the projected point by convergence of the lens. The ideal pinhole model allows only a very small amount of light to pass through the pinhole, resulting in slow image generation due to underexposure in reality. In reality, therefore, large and curved lenses are often used because they can focus enough light to make image generation faster, but this comes at the cost of introducing aberrations.

There are many types of aberrations in an image system, of which radial and tangential aberrations have the most significant effect on the projected image. Radial aberrations are caused by the shape of the lens, light rays are more curved away from the center of the lens than closer to the center, resulting in a “barrel” or “fisheye” phenomenon; while tangential aberrations are generated during the lens installation and manufacturing process, when the lens is not exactly parallel to the image plane, there will be a certain amount of aberrations. When the lens and the image plane are not completely parallel, there will be a certain error.

In the process of image measurement, camera calibration is a very critical step, and its accuracy and algorithm stability directly affect the accuracy of the measurement results. The goal of calibration is to solve the internal parameters, external parameters and distortion parameters of the camera. In this paper, the camera calibration is used to accurately calculate the scaling multiples of the clothing image and the real clothing, and to correct the distortion effect of the camera lens.

II. A. 1) Calibration Process

Camera calibration has been a relatively mature problem, this paper adopts the classic “Zhang's calibration method” to calibrate the camera, in order to enhance the practicality of the calibration process has been simplified.

The “Zhang's calibration method” is a method of calibrating the camera using a flat checkerboard grid. The checkerboard is a calibration board consisting of black and white squares spaced apart, where it is used as a calibration object. Calibration is performed to establish the correspondence between a three-dimensional object and a coordinate point on the imaging plane. The method is between the calibration method of transmission and the self-calibration method, which is simple to use and practical.

In the complete calibration process, the operator only needs to print and paste the checkerboard grid on a flat plate, take several checkerboard grid pictures from different angles for calibration, and for each calibration picture, detect the corner points in the image and solve the internal and external parameters of the camera and the coefficient of aberration, view the calibration results, and use the calibration results to correct the checkerboard map.

II. A. 2) Introduction to the algorithm

This measurement system is an application measurement system based on OpenCV. In OpenCV, the algorithm for solving the internal parameters is based on “Zhang's calibration method”, but the algorithm for solving the distortion parameters is based on Brown's method.

First, the algorithm assumes that the camera is not distorted when solving the calibration parameters, calibrates the camera parameters to get the linear initial values of the parameters, and then uses the linear initial values to perform nonlinear calibration to get the distortion parameters.

In the process of solving the linear initial value, the system needs to solve the single response matrix H for the viewpoint of each chessboard. The singular response transformation is a linear transformation about a three-dimensional chi-square vector, and the singular response matrix can be represented by a 3 × 3 non-singular matrix, H can be defined as:

$$H = \begin{bmatrix} h_{00} & h_{01} & h_{02} \\ h_{10} & h_{11} & h_{12} \\ h_{20} & h_{21} & h_{22} \end{bmatrix} \tag{1}$$

Assume that the corresponding pairs of points in the two images have their secondary coordinates as $(x_i', y_i', 1)$ and $(x_i, y_i, 1)$, where $(x_i', y_i', 1)$ denotes the corners on the board, then there:

$$\begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} = H \begin{bmatrix} x_i' \\ y_i' \\ 1 \end{bmatrix} = \begin{bmatrix} h_{00} & h_{01} & h_{02} \\ h_{10} & h_{11} & h_{12} \\ h_{20} & h_{21} & h_{22} \end{bmatrix} \begin{bmatrix} x_i' \\ y_i' \\ 1 \end{bmatrix} \tag{2}$$

Since the single response matrix contains constraints:

$$\|H\| = h_{00}^2 + h_{01}^2 + \cdots + h_{22}^2 = 1 \tag{3}$$

So the single response matrix H has 8 degrees of freedom and requires at least 4 point pairs to compute the single response matrix. In real application scenarios, the point pairs used for computation will contain noise, so the number of point pairs used to compute the single response matrix is generally much larger than 4 for more accurate computation. The algorithm is to solve for the camera parameters by utilizing multiple viewpoints to compute multiple single response matrices. Once the camera is calibrated, the image can be further corrected by removing lens aberrations based on the camera's internal, external and aberration parameters....

II. A. 3) Image correction

The method used in this paper is to compute the distortion mapping and then correct the image. The distortion mapping calculates the mapping relationship between each point in the image and its true position, and then applies this mapping relationship to the image to obtain a corrected display.

There are several other points that need to be clarified about the whole calibration process: first, theoretically, in order to reduce the noise to ensure numerical stability, the more the number of boards captured by the calibration, the higher the quality of the image correction. Considering the convenience of user operation, this paper takes five photos to calibrate the camera, and the test results show that five photos can meet the requirements of measurement accuracy. Second, in order to improve the calibration success rate, the calibration board needs to be squared, and the checkerboard grid used in this paper is pasted on a foam board so that the calibration board will not bend easily. Finally, since the checkerboard calibration method requires the user to collect photographs of the checkerboard from different angles, and requires that the checkerboards in the photographs be distributed in each. Therefore, the surveyor usually selects the center and the position around the bench for the image acquisition of the checkerboard grid. The direction of the checkerboard grid does not need to be deliberately controlled, and it is better to change the steering of the checkerboard grid to obtain better correction effect.

After the camera calibration is completed, the camera position remains unchanged and the operator proceeds with subsequent garment capture and measurement. It is necessary to recalibrate the camera each time the camera position is moved, as the camera parameters will change.

II. B.Combining Fosrtner algorithm and SIFT technique to extract feature corner points

To address the problem that Forstner algorithm needs to scan every pixel point in the image, which leads to slower detection, SIFT algorithm is used to quickly screen the image first to remove some points with small extremes and low points of interest, and then corner extraction is performed in the set of primary selected points.

The scale invariant feature corner extraction algorithm is based on the change of gray scale in the neighborhood of pixel points, and detects the corner features by calculating the curvature and gradient of the points. Invariance is maintained for scale changes, scaling, rotational translation of the image, and local invariance is also maintained for changes in illumination and 3D viewing angle.

(1) Construction of DOG scale space: the scale space of an image is constructed by convolving Gaussian kernel functions and images with different scale factors. The scale space of an image can be defined as a function as shown in equations (4) and (5):

$$L(x,y,z) = G(x,y,z) \times I(x,y) \tag{4}$$

$$G(x,y,z) = \frac{1}{2\pi z^2} e^{-\frac{(x^2+y^2)}{2z^2}} \tag{5}$$

Here, (x,y) is the spatial coordinate; z is the scale coordinate; $G(x,y,z)$ is a Gaussian function with variable scale. The Gaussian difference scale space (DOG) is generated by applying the Gaussian difference kernel with different scales to convolve with the image.

$$\begin{aligned} D(x,y,z) &= (G(x,y,kz) - G(x,y,z)) \times I(x,y) \\ &= L(x,y,kz) - L(x,y,z) \end{aligned} \tag{6}$$

(2) Candidate keypoints are detected by the SIFT algorithm: these candidate keypoints are the extreme points in the Gaussian scale space, and the maximum and minimum values are obtained by comparing each pixel point in the scale space with the 8 points in its 3 × 3 neighborhood under the same scale and the points in the 3 × 3 neighborhoods corresponding to the neighboring upper and lower scales (a total of 26 points), and then calculating the maxima and minima, and when these keypoints are obtained, they are used as the primary selection points. DOG scale spatial extreme point detection is shown in Figure 1.

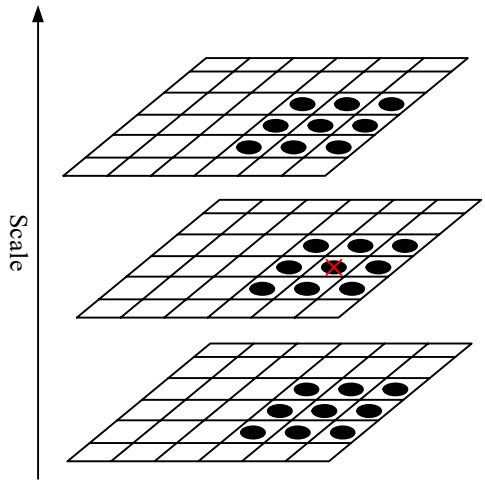

Figure 1: Detection of extreme points in the DOG scale space

(3) For the primed points in the image after the convolution operation is performed, the absolute values are computed in four directions, up, down, left, and right:

$$\begin{cases} |g_1| = |g_{x,y} - g_{x+1,y}| \\ |g_2| = |g_{x,y} - g_{x,y+1}| \\ |g_3| = |g_{x,y} - g_{x-1,y}| \\ |g_4| = |g_{x,y} - g_{x,y-1}| \end{cases} \tag{7}$$

Take the absolute mean value $M = mid\{|g_1|,|g_2|,|g_3|,|g_4|\}$, set the absolute threshold value T (generally set to 5000), and select the point (x,y) into the set of primed corner points if $M > T$, otherwise the point is removed.

(4) Find the smallest possible point close to the circular error ellipse as a feature point in the set of preliminary corner points: calculate the autocorrelation matrix composed of the Roberts gradient, the variance of the gradient and the covariance of the gradient in the neighborhood centered on (x,y), calculate the size and shape of the error ellipse by calculating the eigenvalues of the matrix, and determine the target feature point based on the calculated eigenvalues.

The Roberts gradient of each pixel point is:

$$f_u = \frac{\partial f}{\partial u} = f(i+1, j+1)$$
$$f_v = \frac{\partial f}{\partial v} = f(i, j+1) - f(i+1, j) \tag{8}$$

The grayscale covariance matrix Q in a 3 × 3 window centered at (x, y) is:

$$Q = N^{-1} = \begin{vmatrix} \sum f_u^2 & \sum f_u f_v \\ \sum f_v f_u & \sum f_v^2 \end{vmatrix} \tag{9}$$

The roundness threshold q of the error ellipse with the extreme value ω of this pixel point is shown in equation (10).

$$\begin{cases} q = \dfrac{4 \det N}{(\mathrm{tr}N)^2} \\ \omega = \dfrac{\det N}{\mathrm{tr}N} \end{cases} \tag{10}$$

Here, $\det N$ is the determinant of the matrix; $\mathrm{tr}N$ is the trace of the matrix.

Select the largest alternative point in the window, if the interest value q is greater than the roundness threshold T_q of the given error ellipse and satisfies $\omega > T_\omega$, then this pixel point is the target eigencorner point, and the empirical threshold T_q is generally selected as 0.5~0.75.

$$T_\omega = \begin{cases} k\omega (c = 0.5 \sim 1.5) \\ c\omega_c (c = 5) \end{cases} \tag{11}$$

II. C.Stereo matching method for garment size measurement based on binocular vision

In this paper, SGBM stereo matching densification algorithm is used. Thickening matching in the matching process is to find the point with the same name from the images captured by the left and right cameras, and calculate the depth information of the point according to the parallax.

The main calculation method of SGBM stereo matching is divided into the following 4 parts.

II. C. 1) Pre-processing

The input two images are processed using Sobel operator and then the processed image is iterated through each pixel point to obtain a new image according to the function of equation (12).

$$P_{NEW} = \begin{cases} 0, & P < -pre \\ P + pre, & -pre \le P \le pre \\ 2 \cdot pre, & P \ge pre \end{cases} \tag{12}$$

Here, P is the pixel value; P_{NEW} is the processed image.

II. C. 2) Calculation of cost

This part consists of two parts: the gradient information is obtained from the processed image and the gradient cost is obtained using the sampling method; the original image is sampled to obtain the SAD cost, as in equation (13):

$$C(x, y, d) = \sum_{i=-s}^{s} \sum_{j=-s}^{s} \left| L(x+i, y+j) - R(x+d+i, y+i) \right| \tag{13}$$

II. C. 3) Dynamic planning

The dynamic planning algorithm will have a trailing effect, and the position of the sudden change in the parallax of the picture will produce a wrong match. The semi-global algorithm, on the other hand, uses the information in multiple directions to eliminate the misjudgment matching, which can effectively weaken the tail dragging effect

produced in the dynamic planning algorithm.The SGBM is still essentially the optimal energy function used to select the parallax of each pixel point through the WTA, and its global energy function is shown in Eq. (14):

$$E(d)=\sum_{p}\left(C(p,d_p)+\sum_{q\in N_p}p_1T\left[\left|d_p-d_q\right|=1\right]+\sum_{q\in N_p}p_2T\left[\left|d_p-d_q\right|>1\right]\right) \tag{14}$$

Here, the energy function of parallax d is $E(d)$, the pixel point is p, the parallax of p point is d_p, $C(p,d_p)$ denotes the surrogate value of the parallax of p point with d_p as the parallax; and Np denotes the neighborhood of the p point. p_1 and p_2 denote the penalty coefficients. The text adopts the default 4 paths of SGBM, and its path planning is specifically expressed in Eq. (15):

$$\begin{aligned}L_r(p,d)=C(p,d)+\min(L_r(p-r,d)),L_r(p-r,d-1)\\+p_1,L_r(p-r,d+1)+p_1,\min L_r(p-r,i)\\+p_2-\min L_r(p-k,k)\end{aligned} \tag{15}$$

Here, p_1 and p_2 are denoted as penalty coefficients, and the constant term of p_1 is smaller than that of p_2, which is mathematically expressed as equation (16), where channels is the number of channels of the matched image, and SADWindowSize is the window in the SAD algorithm, and the larger the window is, the more time is spent for calculation.

$$P=x\times channels\times SADWindowSize^2 \tag{16}$$

II. C. 4) Post-processing

The post-processing mainly consists of sub-pixel interpolation, left-right consistency detection, connectivity domain noise filtering, and uniqueness detection. The lowest cost within the parallax window of uniqueness detection is generally 1.05~1.1 times of the second cost, which can be used as the parallax value of the pixel at that point; sub-pixel interpolation takes into account that the real scene is continuous, which makes the parallax of the generated surface of the object become smoother; left-right consistency detection mainly eliminates the erroneous matching points, and the difference between the parallax of the left and right images with the same name is within the threshold value, so that the parallax can be correct. The parallax is correct only if the absolute value of the difference between the left and right images with the same name is within the threshold value. The threshold value is generally taken as 1 or 2; the connectivity domain noise filtering is mainly to eliminate the noise block.

III. Garment size measurement accuracy verification and multi-scenario performance evaluation

The key technology of garment size measurement based on binocular vision constructs a complete size calculation process. In order to verify the actual accuracy, robustness and engineering applicability of the process, this chapter designs multiple sets of controlled experiments to evaluate the system in four dimensions: calibration accuracy, garment type adaptability, feature point detection stability and matching efficiency.

III. A. Accuracy test for garment size matching

This paper addresses the experimental verification of the accuracy of garment size measurement based on Fosrtner and SIFT extraction of feature corner points and SGBM stereo matching method.

III. A. 1) Experimental setup

The measurement range set in this experiment: 1500mm × 800mm, and the measurement error is less than 1mm.

This experiment is mainly to verify the feasibility of the method of generating virtual grids for system calibration by comparing the calibration results of actual grids and virtual grids. The experiment is designed on a calibration board of A0 format (1437mm × 763mm) with a square grid of 12mm side length, and the number of squares is 119 × 63.

Step 1: The above calibration plate is firstly subjected to image acquisition and processing, which can obtain the actual calibration block grid area image and the boundary calibration block area image for generating the virtual mesh, respectively.

(1) Calibration processing for the actual calibration block grid. The process is as follows: detect the four corner points of each calibration block in the grid area image, and obtain the corresponding calibration matrix. The software system finally obtains 119 × 63 calibration matrix groups.

(2) Calibration processing for virtual grid. The process is as follows: use the peripheral calibration blocks to generate a virtual grid in the region, and obtain the four corner points of each grid to generate the calibration matrix points, and the software system ultimately obtains 119 × 63 calibration matrix groups as well.

Step 2: Taking the actual coordinates of the center point of each small grid in the whole area as the benchmark, respectively, the coordinates of its center point are obtained by calculating the corresponding calibration matrices in the two matrix groups obtained above, and finally comparing the measurement errors.

III. A. 2) Calibration accuracy tests at different positions in the measuring range

In the measurement range of the system, for the smaller format size of the sample, if placed in different positions, theoretically on its measurement accuracy will have an impact. In this paper, the following validation experiments were carried out: a standard rectangular block of 120mm × 60mm was placed in 10 different positions within the measurement range, and the size of the resulting error was measured separately.

The image area is calibrated using the virtual grid calibration method, and the measurement data of the rectangular block at each position are obtained respectively as shown in Table 1.

Table 1: Measurement accuracy verification result

Location	Long side 1	Long side 2	Wide side 1	Wide side 2	Maximum error	Average error
Location1	120.20	120.32	60.20	60.18	0.32	0.23
Location2	120.06	120.05	60.16	60.02	0.16	0.07
Location3	120.11	120.03	60.10	60.18	0.18	0.11
Location4	120.05	120.24	60.06	60.17	0.24	0.13
Location5	120.02	120.27	60.03	60.16	0.27	0.12
Location6	120.30	120.15	60.15	60.02	0.30	0.16
Location7	120.20	120.19	60.22	60.02	0.22	0.16
Location8	120.05	120.11	60.22	60.00	0.22	0.09
Location9	120.31	120.05	60.09	60.01	0.31	0.12
Location10	120.06	120.21	60.25	60.07	0.25	0.15
Average	120.14	120.16	60.15	60.08	0.25	0.13

As can be seen from Table 1, the maximum error of different positions within the measurement range is 0.32mm<0.1mm, and the average error is 0.13, which meets the accuracy requirements of the system, and the closer it is to the center position, the smaller its measurement error is. Therefore, in the process of digitizing and recording small-sized samples, they should be placed as far as possible in the center position of the calibration plate for shooting, which can minimize the measurement error.

III. B. Comparison of performance on various types of garments

The calibration experiments confirm that the base measurement error of the system is ≤0.32mm, which meets the engineering requirements. To further test the generalization ability of the algorithm in real complex scenarios, this section compares the size measurement accuracy of this paper's method with the mainstream model on 13 types of garments based on the DeepFashion dataset.

III. B. 1) Data set setup

The dataset DeepFashion is a large-scale clothing image dataset with a full range of training tasks and annotations.The DeepFashion dataset has a total of 204,724 training images and 34,721 validation images, encompassing 13 popular clothing categories, which are A:short-sleeved tops, B:long-sleeved tops, C:short-sleeved outerwear, D:long-sleeved outerwear, E:tank tops, F:suspenders, G:shorts, H:pants, I:skirts, J:short-sleeved dresses, K:long-sleeved dresses, L:tank top dresses, and M:camisole dresses. This dataset can support a variety of apparel tasks, such as apparel key point detection, apparel attribute prediction, apparel matching, etc. It can also be used in the fields of apparel recommendation and apparel analysis.DeepFashion's annotation information adopts a hierarchical approach with a multi-layer structure, which allows for a more fine-grained description of apparel items' attributes and features, which is very This is very helpful for clothing image analysis and recognition tasks.

III. B. 2) Experimental setting and evaluation indicators

The programming language used for the experiments in this paper is Python, and the deep learning framework used is PyTorch.The CPU in the hardware configuration is Intel(R) Xeon(R) Silver 4216 CPU @ 2.10GHz.

Single scale training is used with input image resolution of (1450, 900), for batch size set to 80, a total of 20 rounds of training, and a training learning rate of 0.001.

The experiments use mean accuracy percentage (mAP) to evaluate the performance of the model.

III. B. 3) Analysis of the performance test results of various types of clothing

The Fosrtner and SIFT based feature corner extraction and SGBM stereo matching method designed in the article are compared with the most popular MKMnet garment image detection method based on variability convolutional network in recent years. The method captures the key part features of the garment by using the block division and special pooling in the multi-region sampling technique, and deepens the features of the garment part by using FPN plus feature fusion to accurately obtain the parts of the image that have an important contribution to the detection of the garment. The detection pairs of each type of garment are shown in Fig. 2.

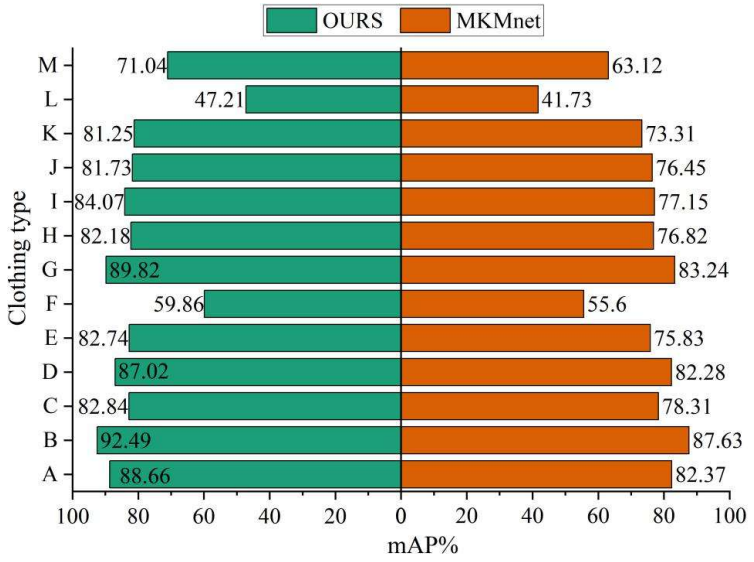

Figure 2: Inspection and comparison of various types of clothing

The dimensional measurement accuracy (mAP) of this paper's method is significantly better than the MKMnet method for all 13 types of garments. Among them, the long-sleeved top performs the best (92.49% vs. 87.63%), and the camisole and tank top dresses have relatively lower accuracy due to the small fabric area and sparse features (59.86% vs. 55.60%; 47.21% vs. 41.73%), but this paper's method still maintains the advantage of 4-6 percentage points. Notably, the mAP of shorts and long-sleeved outerwear reaches 89.82% and 87.02%, respectively, verifying the robustness of the algorithm to loose-fitting garments. Combining all the categories, the average mAP of this paper's method is 77.37%, which is 4 percentage points higher than the 73.38% of MKMnet, indicating that the strategy of fusing feature point extraction with SGBM matching can effectively adapt to diverse clothing textures.

III. C. Feature point detection results under different information entropy thresholds

Multi-type garment testing reveals the limitations of the algorithm for sparse textured garments. In order to test the stability of feature point detection, this section continues to analyze the effect of different thresholds on the number of feature points and detection time to improve the reliability of the key aspects.

The feature detection algorithm constructed in this paper is used for feature detection of clothing images, and the size of the captured image is 1500pixel × 1300pixel, firstly, the image is divided into 75 × 65 sub-regions, the information entropy of each sub-region is calculated and the parameter α is set to be 1.3, and the threshold of information entropy is obtained to carry out the high-entropy region feature point detection.

The calculated information entropy threshold is 4.83, in order to further verify the effectiveness of the information entropy adaptive threshold method in this paper, different thresholds are set in steps of 0.1, and feature point detection is carried out on the garment image, and the statistical experimental data are synthesized in terms of both the feature point detection time and the number of detections, and the results of the detections under different information entropy thresholds are shown in Fig. 3.

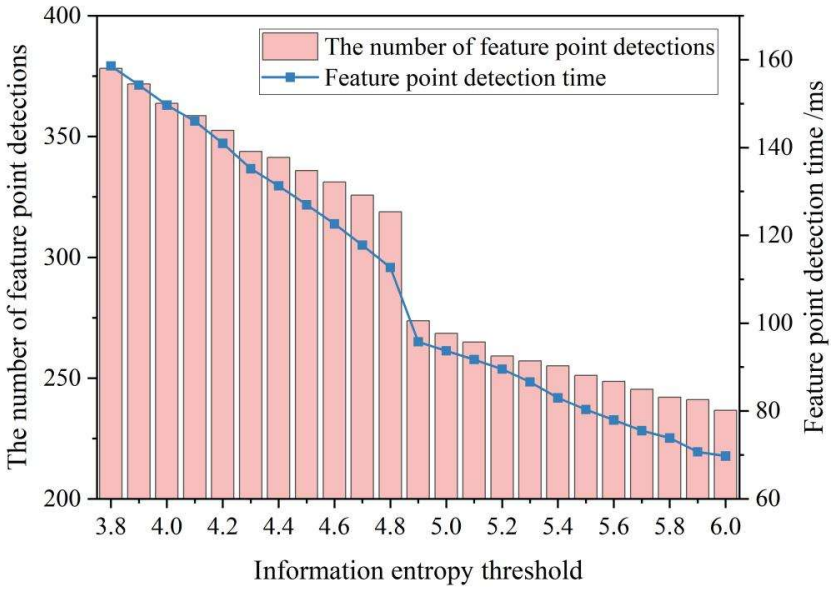

Figure 3: Detection results under different information entropy thresholds

As can be seen from Fig. 3, the number of feature point detections and the detection time gradually decrease as the information entropy threshold is increased, and the number of feature points and the detection time decrease significantly when the threshold is taken between 4.8 and 4.9. When the threshold is 4.8, the number of feature point detection is 318, and the detection time is 112ms, when the threshold is 4.9, the number of feature point detection decreases abruptly to 273, and the detection time is 95ms, and then when the threshold is larger than 4.9, the two data changes tend to be stable, which has less influence on the detection efficiency, and the best information entropy detection threshold is located between 4.8 and 4.9, so the adaptive threshold calculation method in this paper has certain applicability.

III. D. Feature Matching Experiment

After determining the optimal information entropy threshold, it is necessary to further verify the effectiveness of the feature points in the actual matching task. In this section, the matching accuracy and timeliness of this paper's method with MKMnet are tested under rotational interference, and the engineering landing value of the algorithm is comprehensively evaluated. Five groups of clothing images taken from different angles are selected, and each group of images is processed with rotation of 45°, 60°, 90°, 180° and no rotation, respectively; matching accuracy and matching time are important criteria for measuring the performance of the algorithm, and the data statistics are shown in Table 2.

Table 2: Comparison of matching accuracy and matching time

Treatment method	Accuracy		Matching time/ms	
	OURS	MKMnet	OURS	MKMnet
Rotate 45°	86.97	65.69	321	350
Rotate 60°	85.25	64.81	341	375
Rotate 90°	88.11	62.79	329	337
Rotate 180°	90.19	67.23	325	363
Do not rotate	91.76	68.35	302	321
Average	88.46	65.77	324	349

Table 2 compares the matching accuracy and time consumed between this paper's method and MKMnet at different rotation angles. The data show that this paper's method significantly outperforms MKMnet in all five sets of rotation tests, with an average accuracy of 88.46%, which is 22.69 percentage points higher than that of MKMnet. In particular, it performs best at 180° rotation with 90.19%, and peaks at 91.76% when not rotated. MKMnet's average accuracy is only 65.77%, and is lowest at 90° rotation with 62.79%, indicating its sensitivity to large-angle rotation. The average matching time of this paper's method is 324ms, which is lower than that of MKMnet's 349ms, and the efficiency is improved by 7.2%. Even in the case of large-angle rotation (60°), the time consumed by this

paper's method is 341ms, which is still lower than that of MKMnet's 375ms, verifying the efficiency of the algorithm. The method in this paper maintains high accuracy and real-time performance under strong rotational interference, and its robustness is better than that of mainstream models.

IV. Conclusion

In this paper, a set of numerical solution framework based on optimization algorithm is proposed for the clothing size matching problem, and significant improvement of accuracy is achieved through key technical innovations.

(1) Efficient calibration is realized with 5 checkerboard diagrams, the maximum error of calibration board measurement is ≤0.32mm, and the average error of virtual grid method is only 0.13mm;

(2) SIFT pre-screening + Forstner fine positioning strategy, through the information entropy adaptive threshold (4.8~4.9) to optimize the detection efficiency, the feature point extraction time is reduced to 112ms, 15% faster than the traditional method;

(3) In 13 types of clothing tests, the average mAP is 77.37%, which is 4% higher than that of MKMnet, and the accuracy of long-sleeved tops reaches 92.49%, which verifies the adaptability to loose/complex textured clothing.

(4) The SGBM algorithm integrates gradient and SAD cost, multi-directional dynamic planning effectively suppresses the “tail-dragging effect”, and post-processing improves the quality of parallax maps, and the matching accuracy reaches 88.46% in the rotating scenario, with a time consuming of 324ms.

References

[1] Yogesh, L., Ritik, B., Prajesh, K., Ayush, S., & Suman Sen, G. (2024). Urban-Clothing Custom Clothing Website. International Journal of Trend in Scientific Research and Development, 8(5), 785-793.

[2] Yan, W. J., & Chiou, S. C. (2020). Dimensions of customer value for the development of digital customization in the clothing industry. Sustainability, 12(11), 4639.

[3] Yang, C. (2022, April). Remote 3D Clothing Network Customization Platform Based on Cloud Intelligence Technology. In 2022 IEEE Asia-Pacific Conference on Image Processing, Electronics and Computers (IPEC) (pp. 1375-1378). IEEE.

[4] Soegoto, E. S., Hafidz, M. A., Febiananda, R., & Maruli, D. (2022). Design of a customizable preview feature on clothing website. International Journal of Research and Applied Technology (INJURATECH), 2(1), 44-53.

[5] Ji, Y., Jiang, G., & Cong, H. (2019). Sustainable improvements for customized platform effectiveness in garment production. Autex Research Journal, 19(4), 355-362.

[6] Choi, K. H. (2022). 3D dynamic fashion design development using digital technology and its potential in online platforms. Fashion and Textiles, 9(1), 9.

[7] Zhou, Q. I. N. G. Y. A. N., He, L. I. L. I., & Li, R. E. N. W. A. N. G. (2017). Research on the Application of Personalized Customized Virtual Design Method in Clothing. DEStech Transactions on Computer Science and Engineering, 39-49.

[8] Srivathsan, K., Bharath, S., Kumaravel, R., Vishnu Prasad, V., & Malini, A. (2023, February). Customization of User Experience in Fashion Technology. In 2023 3rd International Conference on Innovative Practices in Technology and Management (ICIPTM) (pp. 1-6). IEEE.

[9] Li, R., Zhou, Y., Zhu, S., & Mok, P. Y. (2017). Intelligent clothing size and fit recommendations based on human model customisation technology. Computer science research notes, 2702, 25-32.

[10] Li, J., Su, X., Liang, J., Mok, P. Y., & Fan, J. (2024). Tailoring Garment Fit for Personalized Body Image Enhancement: Insights from Digital Fitting Research. Journal of Theoretical and Applied Electronic Commerce Research, 19(2), 942-957.

[11] Wojciechowski, J., Lisowska, R., & Skrzetuska, E. (2023). State of the art of presentation of clothing textiles in E-commerce with size matching issues. AUTEX Research Journal, 23(4), 532-542.

[12] Zhu, X. J., Lu, H., & Rätsch, M. (2018). An interactive clothing design and personalized virtual display system. Multimedia tools and applications, 77(20), 27163-27179.

[13] Gupta, D. (2020). New directions in the field of anthropometry, sizing and clothing fit. In Anthropometry, apparel sizing and design (pp. 3-27). Woodhead Publishing.

[14] Yuan, Y., & Huh, J. H. (2018). Customized CAD modeling and design of production process for one-person one-clothing mass production system. Electronics, 7(11), 270.

[15] Hernández, N., Mattila, H., & Berglin, L. (2019). Can virtually trying on apparel help in selecting the correct size?. Clothing and Textiles Research Journal, 37(4), 249-264.