3D Snow Sculpture Reconstruction Based on Structured-Light 3D Vision Measurement

Wancun Liu 1,2, Liguo Zhang 3,4,* , Xiaolin Zhang 1 and Lianfu Han 5,*

1 School of Instrumentation Science and Engineering, Harbin Institute of Technology, Harbin 150001, China; Wancun_Liu@163.com (W.L.); zhangxiaolin@hit.edu.cn (X.Z.)
2 School of Mechanical and Electrical Engineering, Harbin Vocational and Technical College, Harbin 150001, China
3 College of Computer Science and Technology, Harbin Engineering University, Harbin 150001, China
4 Heilongjiang Hengxun Technology Co., Ltd., Harbin 150080, China
5 College of Electronics Science, Northeast Petroleum University, Daqing 163318, China
* Correspondence: zhangliguo@hrbeu.edu.cn (L.Z.); Lianfuhan@nepu.edu.cn (L.H.);
Tel.: +86-136-9462-6215 (L.Z.); +86-182-4962-9368 (L.H.)

Abstract: Structured-light technique is an effective method for indoor 3D measurement, but it is hard to obtain ideal results outdoors because of complex illumination interference on sensors. This paper presents a 3D vision measurement method based on digital image processing to improve resistance to noise of measuring systems, which ensures normal operation of a structured-light sensor in the wild without changing its components, and the method is applied in 3D reconstruction of snow sculpture. During image preprocessing, an optimal weight function is designed based on noise classification and minimum entropy, and the color images are transformed into monochromatic value images to eliminate most environmental noise. Then a Decision Tree Model (DTM) in a spatial-temporal context of video sequence is used to extract and track stripe. The model is insensitive to stubborn noise and reflection in the images, and the result of the model after coordinate transformation is a 3D point cloud of the corresponding snow sculpture. In experimental results, the root mean square (RMS) error and mean error are less than 0.722 mm and 0.574 mm respectively, showing that the method can realize real-time, robust and accurate measurement under a complex illumination environment, and can therefore provide technical support for snow sculpture 3D measurement.

Keywords: structured light; 3D measurement; monochromatic value; stripe extraction

1. Introduction

Digitalized archive of cultural relics and artworks has become an important means for their restoration and virtual museum construction thanks to the development of 3D measurement technique. Snow sculpture is an aesthetical form of artwork peculiar to cold regions around the world in winter, as shown in Figure 1, and its cultural value is no less than painting, wood carving, architecture and other artistic expressions [1,2]. However, when warmed up, these artworks, great displays of the wisdom of their designer artists, will disappear. Snow sculpture can be divided into two categories from the perspective of creation [3]. The first category is designed and constructed according to three-dimensional computer modeling, the 3D data of which is known in advance, meaning that there is no need to measure. The other category is works improvised according to shape and texture of snow billet and combined with surrounding scenes. Similar to other sculptures, they fuse the ideas and inspirations of artists. The category of snow sculpture is unique and needs to measure and record 3D data. Therefore, both snow sculpture designers and management companies want to obtain and store 3D data of each snow sculpture in the computer, thereby making it easy to reproduce the original appearance of snow sculpture through technologies such as 3D movie and 3D printing. Obviously, there is actual demand for 3D
measurement of snow sculpture and its scene reconstruction [2], but relevant research has rarely been carried out.

![Figure 1. Snow sculpture artwork.](image)

At present, the main methods for 3D measurement of cultural relics and artworks include laser scanning, sequence image method and structured-light vision measurement [4–7]. Laser scanning is highly accurate but expensive, and its point cloud is not dense enough to adequately describe details. Sequence image method has low requirements for equipment and is easy to operate, but it cannot measure objects without obvious texture (such as pure white porcelain bottles) and has low accuracy. Structured-light vision measurement has the advantages of high accuracy, simple equipment, easy operation and low cost, but the camera is easily affected by illumination. Considering large size, single surface color, complex texture, and low cost of snow sculpture, structured-light vision measurement is more suitable for 3D measurement of snow sculpture.

Structured-light sensors need to work in a relatively controlled light environment traditionally [8], so it is usually operated in a room, while measuring outdoors has always been a challenge. However, snow sculpture is directly exposed to the outdoor environment with sunlight, shadows, reflections, and other environmental noise factors everywhere. Night scanning can reduce the noise, but it is adverse to operator’s health and work quality in the temperature range of −20 °C to −30 °C. 3D laser scanner can perform 3D measurements such as scanning ancient buildings in the day, but its sparse scanning point cloud cannot fully describe the details of snow sculpture. Therefore, we take advantage of a structured-light sensor with a laser as the light source to scan snow sculpture precisely in the field. Moreover, compared to the coding mode, which is mostly used in controlled scenes, the fixed mode structured-light sensor with better anti-noise performance [9] is more applicable for 3D measurement of snow sculpture.

It is hard to obtain a higher signal to noise ratio (SNR) even when using a laser projection outdoors, because the intensity of sunlight is usually 2–5 orders of magnitude higher than that of structured-light stripe [8]. The existing methods have not well resolved this problem. For example, Microsoft’s Kinect operates difficulty in high ambient light [10]. Modifying the structured-light sensor, such as installing filter or increasing the intensity of the light source, is a method to reduce noise, but it is at the expense of robustness and cost. So many researchers prefer to solve the above problems by digital image processing.

Stripe extraction is the key factor for fixed mode structured-light equipment to improve measurement accuracy, because the calibration and coding processes have been fixed. In Reference [8], Gupta et al. proposed an adaptive stripe intensity method to improve SNR by adjusting the width of stripe according to the intensity of sunlight, which does not increase the extra power consumption, but reduces the scanning efficiency. In Reference [11], O’Toole et al. analyzed the stripe collected by a camera on direct and indirect light projection paths, and the ambient noise was directly ignored, but the optical simulating model is too complicated. In Reference [12], Steger proposed an unbiased detection method based on Hessian matrix, which can locate the curve points in sub-pixel accuracy and has strong anti-interference ability, but it relies heavily on segmentation results of stripe area. In Reference [13], Usamentiaga et al. segmented foreground information using center of
Then the relative motion of two trapezoidal windows was used to search the laser skeleton, and curve fitting was used to mend the laser stripe gap. All the above methods focus on eliminating noise and extracting light stripe from a single image, while few establish frame-to-frame correlation of consecutive video. Inter-frame correlation can be regarded as a process of target tracking. In the presence of environmental noise, surface reflection and partial occlusion, a combination of spatial and temporal information can be used to detect laser stripe stably and rapidly. In Reference [14], Janne used Kalman filter to track segmented light stripe, and their brightness, length, direction and position are used to establish the correlation of adjacent frames. This method is used to detect obstacles outdoors, but it is only suitable for part of outdoor scenes.

To solve the problem that a structured-light sensor is disturbed heavily by environment noise in snow sculpture measurement, a 3D vision measurement method focused on light stripe extraction and tracking is proposed in this paper, as shown in Figure 2.

The original images mixed with noise collected by the structured-light sensor are classified based on RGB color space distribution and histogram image at first. Then optimal weight combination of RGB is constructed, and the color image is transformed into a monochromatic value image with high SNR. At last, a Decision Tree Model is established in a spatial and temporal context (STC-DTM) to extract and track the laser stripe. The result of the model after coordinate system transformation is a 3D measurement point cloud. This method improves the performance of a measurement system through image processing technique without changing the components of the structured-light sensor. It is universal for other similar 3D measurements.

The remainder of this paper is organized as follows. Section 2 provides details of the method and their algorithms. Section 3 describes the result of experiments for evaluating the accuracy, stability and speed of the method. The final section discusses and concludes the paper.
2. Methods

2.1. Noise Classification

The laser stripe in a structured-light image is the foreground, while othersignals interfering with the stripe extraction are noise. By analyzing snow sculpture measurement in actual scene, we found the main noise is from sunlight, shadows, surface color, surface reflections, and occasionally colored lights. As multiple types of noise are superimposed on images, it is hard to analyze such multi-source noise by constructing an optical path model.

In space distribution, light and surface color are global noise, whereas shadow, surface texture, and reflections are local. Snow sculpture is generally located in an open field, surrounded by snow. Since there are generally no colored buildings, flowers or other objects around, the possibility of color pollution can be virtually eliminated. During measurement, the structured-light scanner is calibrated with sunlight. According to color distribution, sunlight and shadow usually have similar R, G, and B components, whereas the color values of colored light and surface colors are usually higher in one color dimension. Reflective noise is related to illumination and surface characteristics, and shows localized high intensity. The laser stripe is overlaid by a large amount of environmental noise under sunlight, and the intensity distribution in R, G and B space is very close, thus presenting a single peak in the histogram, as shown in Figure 3. Conversely, the similarity between the three histograms is greatly reduced when there is interference from colored light or surface color. In the case of surface reflection, the similarity of intensity in space distribution decreases.

Based on the above analysis of color histogram, a qualitative description of the ambient noise in R, G and B space is made, which represents an approximate Gaussian distribution with different mean values and variances, and it will guide the color value space transformation.

2.2. Monochromatic Value Space Transformation

RGB color space sometimes does not show part of information we need, and some important information cannot be distinguished from the color space. Therefore, transforming a color value space to a monochromatic value space is more conducive to the extraction of laser stripe. In this paper, a linear transformation is used because it does not cause large jumps, singular value points, or discontinuities, and is faster than nonlinear transformation [15]. A gray image is the most typical monochromatic value image generated by linear transformation.

In a discrete color snow sculpture image \( f_{ij} \) (with a size of \( M \times N \)), the tristimulus values of the pixels are \( R_{ij}, G_{ij} \) and \( B_{ij} \). The linear transformation is defined as Equation (1).

\[
F_{ij} = w_r R_{ij} + w_g G_{ij} + w_b B_{ij}
\]  

(1)

where \( F_{ij} \) is the monochromatic value image after transformation, and \( i \) and \( j \) represent the row and column indexes of pixels respectively, with \( i = 1, 2 \ldots, M, j = 1, 2 \ldots, N \), \( w_r, w_g, w_b \in R \). From Equation (1), the monochromatic value image is determined by
the transformation coefficients. So, we need to define an objective function to search for the optimal \( w_r, w_b \) and \( w_j \) to highlight the laser characteristics.

The projected stripe of the laser is relatively concentrated and presents a concise structure, and its profiles showing all the luminance row vectors are shown in Figure 4. There is a high contrast between the laser and background, implying a high Signal-to-Noise Ratio (SNR). The objective function should make full use of this characteristic so that the laser stripe can be extracted easily after transformation. Therefore, the objective function is taken as a measure of the contrast.

![Figure 4](image.png)

**Figure 4.** Profiles of the laser. (a) Images captured including laser stripe on snow sculpture surface; (b) Luminance values row by row.

The pixel values in the laser stripe are close to their averages in Figure 4. In these areas, the contrast increases with energy concentration, and the waveform is steep, suggesting that kurtosis is high. This can be seen as a contrast between the laser stripe and the background. Thus, it is reasonable to choose the kurtosis to define the contrast. The kurtosis \( K \) is defined as Equation (2).

\[
K = \frac{k_4}{k_2^2} = \frac{\mu_4}{\sigma^4} - 3, \tag{2}
\]

where \( k_4 \) is the fourth cumulant, \( k_2 \) is the second cumulant, and \( \mu_4 \) and \( \sigma^4 \) are the mean of fourth moment and the square of the variance \( \sigma \) of the probability distribution respectively. In particular, the kurtosis is zero in the case of normal distribution.

We expect that the laser stripe has the maximum kurtosis after transformation, while the noise signals from background have very low kurtosis, suggesting they are very disordered. In signal systems, disorder is synonymous with entropy. Entropy is affected by the amount of information and its random properties. The more information there is, the more disordered the signals are, and the greater the entropy is. Wiggins [16] first used the minimum entropy deconvolution technique and called it the minimum entropy method. Therefore, maximizing kurtosis is equivalent to minimizing entropy, and the transformation model can be called the minimum entropy model. When \( K \) is less than 0 or if the signals come from a pulse source, minimizing is better than maximizing for the purpose of this study [17,18]. Therefore, the objective function can be further defined as Equation (3).

\[
\text{abs}(K) = \left| \frac{\mu_4}{\sigma^4} - 3 \right|. \tag{3}
\]

Equation (3) is fully differentiable, thus the maximizing function can be defined as Equation (4).

\[
K^2 = \left( \frac{\mu_4}{\sigma^4} - 3 \right)^2. \tag{4}
\]
The monochromatic value image $F_{ij}$ obtained after transformation is regarded as a multi-channel signal (with $N$ segments and $M$ elements per segment), and the Kurtosis can be written as

$$E = K^2 = \left( \frac{\sum_{j=1}^{N} (F_{ij} - \mu_j)^4}{\left( \sum_{j=1}^{N} (F_{ij} - \mu_j)^2 \right)^{\frac{3}{2}}} - 3 \right)^2, \quad (5)$$

where $\mu_j$ is the mean of column $j$ in the image $F_{ij}$.

The solution of maximizing (5) is given in Equation (6):

$$\frac{\partial E}{\partial w_r} = \frac{\partial E}{\partial w_g} = \frac{\partial E}{\partial w_b} = 0, \quad (6)$$

Obviously, it is difficult to solve the above equation. We can learn such knowledge from Robert t. Collins et al. [19]. The continuous coefficient $w_r, w_g$ and $w_b$ in Equation (1) determine an infinite color feature set. To calculate expediently, $w_r, w_g$ and $w_b$ are discretized as integer confined in $[-2, 2]$, and some common color spaces and models are also covered by this range, such as R + G + B, R−B, and so on, implying that it is feasible.

The laser used in the paper is red, and R component of the original image is higher in content than the other components, so limit $w_r \geq 0$. The value ranges of $w_r, w_g$ and $w_b$ are $w_r \in \{0, 1, 2\}, w_g, w_b \in \{-2, -1, 0, 1, 2\}$, where $w_r, w_g$ and $w_b$ are not allowed to be zero at the same time. Then the coefficient vector can be solved by the traversal method. In Figure 5, a red single line laser is used. The optimal coefficient vector of the captured image is $(w_r, w_g, w_b) = (1, -1, 0)$, and the monochromatic value space is R−G, whose transformation result of laser stripe is shown in Figure 5d. For comparison, Figure 5c shows the result in typical gray scale space $(w_r, w_g, w_b) = (0.30, 0.59, 0.11)$. It is obvious that R−G space has more advantages in noise suppression and better SNR is obtained from it. The laser stripe is clearly distinguished from the background in R−G space, and the amount of data is only one third of the original image in Figure 5a.

![Figure 5](image_url)

**Figure 5.** The result of color feature transformation. (a) Original image; (b) R component image; (c) Grayscale image; (d) Monochromatic value of R−G image.

To reduce harm of cold environment to people, we usually measure snow sculpture in the daytime. Also, snowy days are not chosen to avoid uneven distribution of new snow and its impact on measurement results. The $R, G, B$ weights in the algorithm of this paper are dynamically changed. During each measurement, we readjust the parameters according to the calculation results to adapt to the measurement environment at that time.

### 2.3. Stripe Extraction and Tracking

After image preprocessing, global noise such as illumination is eliminated to the maximum, and the geometry edge of the laser stripe is kept continuous. However, some high light regions caused by strong reflection of snow sculpture surface are similar to laser projection, and the laser stripe presents different deformations along the textured target surface, which can cause a large deviation during laser stripe extraction and tracking. With traditional methods it is difficult to balance robustness, accuracy and real-time performance.
In this paper, a Decision Tree Model based on spatial and temporal context (STC-DTM) is proposed to solve these problems.

2.3.1. Establishment of a Spatial Decision Tree Model (S-DTM)

In a single frame, there are two basic spatial constraints between the laser stripe and background nearby: continuity and uniqueness [20]. Although the laser stripe is a smooth and continuous line, the jump or break points indicate it is shaded. In addition, in the same frame, the laser stripe cannot appear in more than one place (parallel or grid light stripe as a whole). Based on these constraints, the image pixels can be classified as two states: on or off the laser stripe. Therefore, stripe extraction can be considered a dichotomy of decision trees, taking the pixels in each column of the image as a sample set, and then being classified by selected features column by column.

The brightness distribution of a laser stripe is shown in Figure 6. We can learn from the figure that the central point is much brighter than neighboring points, and the distribution is continuous. So, the brightness of central points is very close between adjacent columns, and their distance is usually very small. Therefore, the brightness difference and the distance are candidate features for our study.

![Figure 6. Brightness distribution of a laser line.](image)

The laser stripe represents foreground information in the image, and it can be segmented effectively by the center of mass [14,21], which can be used to thin the light stripe. Canny edge detection is performed on the obtained monochromatic image to find edge pixel points, and match adjacent edge pixel points in each column to form point-pairs. In this way, the laser stripe presents a series of point-pairs. The distance between point-pairs is relatively small and their positions are concentrated. Moreover, the brightness in the point-pairs is also very similar to the corresponding monochromatic value image. So, we can calculate the center-of-mass position of R component for each point-pair by Equation (7).

\[
M_r(w) = \frac{\sum_{j=0}^{m} I_{ij} \times f}{\sum_{j=0}^{m} I_{ij} + \sum_{j=0}^{m} f}, \quad i \in [0, n], \quad j \in [0, m),
\]

where \( w \) is window value (in the unit of pixel) of the calculated area, \( I_{ij} \) is brightness value of the pixel at row \( i \) and column \( j \) on the image, and \( M_r(w) \) is column coordinates. After obtaining a series of brightness peaks, the laser stripe is refined into a chain of laser skeleton, as shown in Figure 7a. However, because of noise, there cannot be only one center-of-mass point in each column, and the light stripe skeleton must be accompanied by some noise points, so we must classify it further.

All the center-of-mass positions obtained in each column are taken as a position set of potential light stripe center points, and each position is represented as \( (x_r(t), t) \), \( t = 1, 2 \ldots T \), where \( T \) is step(column) number of the image, \( r \) is the index at step \( t \), and \( r \in [1, R(t)] \), \( R(t) \) is the amount of center-of-mass points at step \( t \), as shown in Figure 7b. The coordinates of the edge pixel point on step \( t \) are \( (l, c_r(t)) \). The mean R component of all the pixels in the point-pair \( ([l, c_r(t)], [l, c_{r+1}(t)]) \) is \( m_r(t) \), and then the brightness difference of the center-of-mass points between adjacent columns can be described as Equation (8).

\[
f_{ij} = |m_r(t-1) - m_r(t)|,
\]

where \( i \) is the \( i \)th point at step \( (t-1) \), and \( j \) is the \( j \)th point at step \( t \).
In addition, the distance between the centers of mass is also an important feature to judge whether they are the center points of the laser stripe, because the center points on the laser stripe do not change position suddenly. The Euclidean distance between two pixel points can be expressed as

$$d_{ij} = \sqrt{(c_i(t - 1) - c_j(t))^2 + 1},$$

where $c_i(t - 1)$ and $c_j(t)$ are column coordinates of the two center-of-mass points.

The selected features are substantiated with data to derive $H_f$ and $H_d$, which are the upper limit values of $f_{ij}$ and $d_{ij}$ respectively, with $\text{Gain}(f_{ij}) > \text{Gain}(d_{ij})$. So, the position set is classified by brightness difference $f_{ij}$ at first, and then estimated by distance $d_{ij}$. The process of the laser stripe extraction by S-DTM is shown in Figure 7b. The result in Figure 7c shows the laser stripe is smoother and clearer, and noise has been well eliminated.

![Figure 7](image_url)

**Figure 7.** The process of the stripe extraction by S-DTM, with “□” being the edge pixel point $[r, c(t)]$ and “■” the center-of-mass point $i_j(t)$. (a) Thinning result by center-of-mass method; (b) Explanation of classification of S-DTM by taking three columns of pixels from the red line in (a); (c) The result of stripe extraction by S-DTM.

### 2.3.2. Establishment of a Temporal Decision Tree Model (T-DTM)

Similar to spatial context correlation, there is a strong temporal relationship between a light stripe and its background in video frame sequences. A traditional research about target tracking proposes that the local context of a current frame helps predict the location of light stripe in the next frame, because the shape, color, position of the target will not change much between adjacent frames, and the rate of change is relatively stable [22,23]. Video sequences captured by a structured-light sensor also have the same characteristics. Ideally, laser lines in adjacent frames are represented as many line segments and arcs with similar shape, area and color, unless they are changed by the texture and reflection of snow sculpture surface, sensor movement and so on. Therefore, the line segments and arcs in each frame may or may not be on the laser stripe, and they also can be classified by a decision tree.

Several possible laser strip regions, which include the optimal solution and several suboptimal solutions of the S-DTM, are defined as a series of sub-windows $W_m(k) = (x, y, I, \Delta, s)$, where $x$ and $y$ are size of the window, $I$ is mean brightness of the laser stripe in the window, $\Delta$ is the change rate of tracking result relative to the previous frame, $s$ is the scale factor, $s \in [0.1, 10]$, $n$ is the amount of candidate windows per frame, $n \in [1, N(k)]$, $N(k)$ is the amount of sub-windows at step $k$, $k = 1, 2, \ldots, F$, and $F$ is the number of the frame. Each sub-window $W_m(k)$ is defined as a position set of the current frame as shown in Figure 8, and measurable features between two adjacent frames are expressed as Equations (10) and (11).

$$P_{mn} = \sqrt{(x_m(k) - x_n(k - 1))^2 + (y_m(k) - y_n(k - 1))^2},$$

$$I_{mn} = |I_m(k) - I_n(k - 1)|.$$
where \( P_{mn} \) and \( I_{mn} \) present the position and brightness change between \( W_m(k) \) and \( W_n(k-1) \) respectively. In sub-window \( W_m(k) \), the average coordinates of all pixels on the laser stripe are \((x_n(k), y_n(k))\) and the mean brightness of R component is \( I_n(k) \).

\[
P_{mn} = \sqrt{(\bar{x}_m(k) - (x_n(k-1) + \int_0^\tau d\mu))^2 + (\bar{y}_m(k) - (y_n(k-1) + \int_0^\tau d\nu))^2},
\]

(12)

where \( d\mu \) and \( d\nu \) represent the velocity of the scanner in the \( u \) and \( v \) directions of the image space respectively, and \( \tau \) is the interval between adjacent frames. So, the change rate of tracking position \( P_\Delta \) is described as Equation (13).

\[
P_\Delta = P_{mn}(k) - P_{mn}(k-1),
\]

(13)

The selected features are substantiated with data to derive \( H_I \) and \( H_{I_\Delta} \), which are the upper limit values of \( I_{mn} \) and \( P_\Delta \) respectively, with \( \text{Gain}(P_\Delta) > \text{Gain}(I_{mn}) \). The sub-window position set is then classified according to the gain values in proper order. The process of the laser stripe tracking by T-DTM is shown in Figure 8, where the filled circles represent the laser stripe positions.

At this point, the STC-DTM has been established. Here S-DTM and T-DTM are not in a cascade relation, but unified as a whole. The video collected by structured-light sensor is the input of STC-DTM, and the output is a series of smooth centerline tracks, which are also a global optimal solution of the model. Although the input of the temporal model depends on the output of the spatial model, optimal solution of the whole model depends on the original data set and the feature set. The optimal solution of the spatial model may be suboptimal in the temporal model, and vice versa.

According to the inherent calibration equation of the structured-light sensor, the tracks from STC-DTM are transformed from an image coordinate system to a sensor coordinate system, and then transformed to a global coordinate system [24] to finally obtain the point cloud data of the snow sculpture surface.

3. Results

To evaluate the validity and applicability of the method proposed in this paper, some experiments have been carried out. In the experiments, the algorithms are implemented in C++ and tested with an Intel Core i7-4790CPU of 3.6 GHz. In the structured-light device, a
red single line laser projector of 650 nm and a color camera of SONY 1/4-in Charge coupled device (CCD) are used, and the angle between them is fixed at 45°. The video capture speed is 25 frames per second, and the image size is 640 × 480 pixels, 5.6 µm × 5.6 µm/pixel, with 8 mm of focal length, and the laser line width is about 2 mm.

3.1. Accuracy Test

In this section, a standard cube, a snow sculpture and a conventional object are measured, and then mean error and root mean square (RMS) error are calculated, respectively. The snow sculpture is large without a standard size, so the value measured by high-precision instrument is taken as the standard value, and the value measured by the method proposed in this paper is taken as the measurement value. The feature distance is used to evaluate measurement accuracy, to avoid the error caused by transformation of point coordinates. The standard value is \(d\), and the measurement value is \(d'\). Then the mean error can be expressed as Equation (14).

\[
\Delta d = \frac{1}{n} \sum_{i=1}^{n} (d'_i - d_i),
\]

where \(n\) is the number of the selected feature line segments, with \(i = 1, 2, \ldots, n\).

3.1.1. Standard Cube Measurement

In the experiment, a standard cube with black and white grid is designed. The size of the cube is 200 mm × 200 mm × 200 mm, the mesh of which is 20 mm × 20 mm, and the uncertainty is 0.01 mm. Six positions of the cube are chosen randomly, where 13 feature distances on the laser line are selected, as shown in Figure 9. The standard values are measured by the vernier caliper (0.02 mm), and the result is shown in Table 1, where \((d_x, d_y, d_z)\) and \((d'_x, d'_y, d'_z)\) represent the values in X, Y and Z directions of the standard value \(d\) and measurement value \(d'\) respectively in global coordinates.

![Figure 9. Standard cube measurement; (a) 13 feature distances selected on the standard cube; (b) 13 feature distances on the 3D point cloud.](image)

The RMS error of the feature distance is 0.125, and \((0.049, 0.147, 0.113)\) in X, Y and Z directions. The mean error of the feature distance is 0.046 mm, and the maximum errors in X, Y and Z directions are 0.099 mm, 0.409 mm and 0.313 mm, respectively.
Table 1. The result of standard cube measurement.

| d_x (mm) | d_y (mm) | d_z (mm) | d' (mm) | d'_x (mm) | d'_y (mm) | d'_z (mm) | Δd_x (mm) | Δd_y (mm) | Δd_z (mm) | Δd (mm) |
|----------|----------|----------|---------|-----------|-----------|-----------|------------|------------|------------|---------|
| 1        | 0.057    | 52.398   | 62.419  | 81.497    | 0.013     | 52.479    | 62.464     | 81.583     | -0.044     | 0.081   |
| 2        | 66.200   | 66.276   | 0.264   | 93.675    | 66.180    | 0.331     | 93.593     | -0.020     | -0.096    | 0.067   |
| 3        | 51.145   | 1.956    | 58.665  | 77.854    | 51.202    | 1.880     | 78.536     | -0.057     | -0.076    | -0.129  |
| 4        | 0.273    | 118.477  | 85.960  | 146.376   | 0.204     | 118.576   | 85.975     | 146.627    | -0.069    | 0.299   |
| 5        | 119.601  | 2.100    | 87.924  | 148.457   | 119.700   | 2.111     | 87.905     | 148.526    | 0.099     | 0.011   |
| 6        | 0.289    | 57.164   | 119.494 | 132.428   | 0.231     | 57.158    | 119.615    | 132.570    | -0.058    | -0.006  |
| 7        | 116.334  | 60.623   | 0.097   | 131.182   | 116.365   | 0.088     | 131.207    | 0.031      | -0.006    | -0.009  |
| 8        | 44.990   | 118.157  | 0.348   | 126.433   | 44.984    | 0.411     | 126.428    | -0.006     | -0.003    | 0.063   |
| 9        | 62.877   | 2.721    | 120.578 | 136.015   | 62.825    | 2.740     | 120.537    | 135.955    | -0.052    | 0.019   |
| 10       | 0.200    | 69.578   | 123.365 | 141.634   | 0.177     | 69.562    | 123.323    | 141.589    | -0.023    | -0.016  |
| 11       | 122.311  | 7.598    | 0.544   | 122.548   | 122.284   | 7.550     | 0.503      | 122.518    | -0.027    | -0.048  |
| 12       | 0.005    | 117.543  | 0.199   | 117.543   | 0.030     | 117.528   | 0.135      | 117.528    | 0.025     | -0.015  |
| 13       | 0.025    | 1.913    | 116.542 | 116.358   | 0.040     | 2.322     | 116.855    | 116.878    | 0.015     | 0.409   |

Mean error (mm): -0.006  0.043  0.025  0.046
RMS errors (mm): 0.049  0.147  0.113  0.125

3.1.2. Snow Sculpture Measurement

A snow sculpture is selected with the approximate size of 3000 mm × 3000 mm × 3000 mm. Twenty feature points are selected on the surface of the snow sculpture, and 20 blue round marking patches with a diameter of about 2 mm and a thickness of 0.1 mm are attached to these points, as shown in Figure 10. At first, a laser tracker is used to measure each pair of feature points for three times, and an average is derived. The distance between two points is then calculated as the standard value. Next, the center of corresponding blue marking patches are in the point cloud generated by the structured-light device. The distance between them is obtained. The average of three measurements is taken as the final measurement. The results are shown in Table 2.

Figure 10. Snow sculpture measurement: (a) 10 feature distances selected on the snow sculpture; (b) 10 feature distances on the 3D point cloud.
Table 2. The result of snow sculpture measurement.

| Standard Value | Measurement Value | Error |
|----------------|-------------------|-------|
| $d_x$ (mm)     | $d_y$ (mm)        | $d_z$ (mm) |
| $d'_x$ (mm)    | $d'_y$ (mm)       | $d'_z$ (mm) |
| $d''_x$ (mm)   | $d''_y$ (mm)      | $d''_z$ (mm) |
| $\Delta d_x$ (mm) | $\Delta d_y$ (mm) | $\Delta d_z$ (mm) | $\Delta d$ (mm) |
|----------------|-------------------|-------|
| 1              |                  |       |
| 2              |                  |       |
| 3              |                  |       |
| 4              |                  |       |
| 5              |                  |       |
| 6              |                  |       |
| 7              |                  |       |
| 8              |                  |       |
| 9              |                  |       |
| 10             |                  |       |

Mean RMS errors (mm)
- 0.700
- 0.558
- 0.178
- 0.302

The RMS error of the feature distance is 0.722, and (0.755, 0.588, 0.862) in X, Y and Z directions. The mean error of the feature distance is 0.574 mm, and the maximum errors in X, Y and Z directions are 1.522 mm, 1.033 mm and 1.409 mm, respectively.

3.1.3. Conventional Object Measurement

In the above two experiments, the measurement accuracy is quite different, so the third object is tested. The third measured object, as a conventional object, is irregular, opaque and less reflective, with a size of 4000 mm × 1000 mm × 2000 mm. 10 feature distances are selected as shown in Figure 11. The measurements are processed in the same way as that in Section 3.1.2, and the result is shown in Table 3.

Table 3. The result of conventional object measurement.

| Standard Value | Measurement Value | Error |
|----------------|-------------------|-------|
| $d_x$ (mm)     | $d_y$ (mm)        | $d_z$ (mm) |
| $d'_x$ (mm)    | $d'_y$ (mm)       | $d'_z$ (mm) |
| $d''_x$ (mm)   | $d''_y$ (mm)      | $d''_z$ (mm) |
| $\Delta d_x$ (mm) | $\Delta d_y$ (mm) | $\Delta d_z$ (mm) | $\Delta d$ (mm) |
|----------------|-------------------|-------|
| 1              |                  |       |
| 2              |                  |       |
| 3              |                  |       |
| 4              |                  |       |
| 5              |                  |       |
| 6              |                  |       |
| 7              |                  |       |
| 8              |                  |       |
| 9              |                  |       |
| 10             |                  |       |

Mean RMS errors (mm)
- 0.539
- 0.297
- 0.435
- 0.302
The RMS error of the feature distance is 0.508, and (0.539, 0.297, 0.435) in X, Y and Z directions. The mean error of the feature distance is 0.159 mm, and the maximum errors in X, Y and Z directions are 0.782 mm, 0.402 mm and 0.625 mm, respectively.

3.2. Speed and Robustness Evaluation

In the present study, two snow sculpture scenarios are selected, corresponding to the lighting conditions at 7:00 a.m. (with little sunlight) and 10:00 a.m. (with much sunlight) respectively. By using structure from motion method with patch-based multi-view stereopsis method processing (PMVS-SFM) brought forward by Furukawa et al. [25], laser scanning method using Leica RTC360 and the method proposed in this paper, 3D reconstruction is performed, with the results given in Figures 12 and 13, respectively.

Figure 12. Three-dimensional reconstruction of snow sculptures at 7:00 a.m. (a) Snow sculpture front; (b) 3D reconstruction results using PMVS-SFM; (c) 3D reconstruction results using laser scanning method; (d) 3D reconstruction results using our method.

Figure 13. Three-dimensional reconstruction of snow sculptures at 10:00 a.m. (a) Snow sculptures side; (b) 3D reconstruction results using PMVS-SFM; (c) 3D reconstruction results using laser scanning method; (d) 3D reconstruction results using our method.

It can be seen from the test results that for snow sculpture with complicated surface texture, a dense 3D point cloud generated with PMVS-SFM image reconstruction method of Furukawa contains more invalid areas, preventing it from well reproducing the original appearance of the snow sculpture. The measurement grows even poorer with the increase of exposure to sunlight. Laser scanning method is less sensitive to sunlight, but provides a relatively sparse point cloud with vague details. Moreover, its laser measurement accuracy of surface texture of the snow sculpture is affected by secondary reflection characteristics of the surface. By comparison, the method proposed in this paper successfully addresses the impact of sunlight and secondary reflection of snow, thus leading to more satisfactory imaging. The benefits include clearer texture and higher stability and accuracy.

Meanwhile, the average execution time of 1000 frames is calculated to be 21.3 ms. It is enough to ensure real-time nature of the normal video capture operation at a speed of 25 frame/s.
4. Discussion

The measurement error statistical result of the three tests in Section 3.1 are shown in Figure 14. The errors are all distributed in a certain range, and the overall measurement accuracy of snow sculpture is lower than that of the other two measurement objects, with large local errors. The reason is analyzed as follows.

4.1. Light Environment and Optical Characteristics of the Measured Object Surfaces Affect the Measurement Accuracy

The surface characteristics of the measured objects include color, texture, brightness, roughness, etc., which produce different noise signals on the collected images. In the outdoor environment, the snow sculpture has strong reflection, complex texture, shadow and other factors that affect the measurement accuracy, which are different from the surface characteristics of standard cube and conventional object, leading to reduced accuracy. It is shown in Table 4 that the multiple additional noise sources lead to the decrease of accuracy and speed, but the measurement error is acceptable for large snow sculpture. It is also proved through the tests that the method is applicable in outdoor complex light environment, meaning that it can provide technical support for digital archiving of snow sculptures.

Figure 14. The measurement errors and their components in x, y and z directions. (a) Error distribution of standard cube measurement; (b) Error distribution of snow sculpture measurement; (c) Error distribution of conventional object measurement; (d) Mean error comparison of three measurements.

4.2. Large Local Error is Related to Complex Texture and the Length of the Feature Distance
Table 4. The relation between the noise and the results in the experiments.

| Measured Object          | Index   | Reflection | Shadow | Surface Color/ Color Light | Mean Error (mm) | Speed (ms/Frame) |
|--------------------------|---------|------------|--------|----------------------------|-----------------|------------------|
| Standard cube(wood)      | weak    | none       | weak   | 0.046                      | 18.9            |
| Conventional object(stone) | middle | strong     | strong | 0.159                      | 20.7            |
| Snow sculpture(snow)     | very strong | very strong | very strong | 0.574                  | 21.3            |

4.2. Large Local Error Is Related to Complex Texture and the Length of the Feature Distance

Some large local errors appear at feature distances where the texture is complex, such as No.9 on snow sculpture, No.10 on conventional object, because serious distortion and occlusion of laser stripe are caused by concave–convex surfaces and lead to inaccurate stripe extraction. In addition, some long feature distances have large errors, because the accumulated errors increase during the measurement without mark points. Therefore, high-accuracy splicing method without mark points is a main task in the next-stage research to improve the accuracy.

5. Conclusions

This paper proposes an accurate, fast and robust 3D measurement method based on structured-light used in complexity light environment. First, an optimal monochromatic value space based on a minimum entropy model is selected for segmentation to maximize the elimination of global noise. Then, a Spatial and Temporal Context Decision Tree Model (STC-DTM) is constructed to extract and track the laser stripe accurately to obtain an accurate and dense 3D point cloud. Finally, the experiments show that this method is effective and applicable to CCD cameras without optical filtering function. Moreover, the method is universal for 3D field measurement and reconstruction in the sectors of industrial detection, cultural archaeology and criminal investigation, and therefore has a good application prospect.

Author Contributions: Conceived the method and edited the manuscript, W.L.; investigation and wrote the paper X.Z.; conceived the method and designed the experiments W.L. and L.H.; project administration and performed the experiments, L.Z. and W.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Science Foundation of China (61501132), National Natural Science Foundation of Heilongjiang Province (LH2020E012) and the China Postdoctoral Science Foundation (2016M601399).

Acknowledgments: The authors would like to thank the associate editor and the reviewers for their helpful comments, which improved the quality of this paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Dewar, K.; Meyer, D.; Li, W.M. Lanterns of ice, sculptures of snow. *Tour. Manag.* 2001, 22, 523–532. [CrossRef]
2. Huang, Z.; Zhang, L. Analysis of and Research into Foreign Factors That Drive China’s Ice-Snow Tourism. In *Communications in Computer and Information Science, Proceedings of the Advances in Applied Economics, Business and Development*, Dalian, China, 6–7 August 2011; Springer: Berlin/Heidelberg, Germany, 2011; pp. 422–427.
3. Li, Y.; Chen, Z.Y.; Wang, C.M. Research on the innovation and development of Harbin ice and snow tourism industry under the background of Beijing winter Olympic Games. *China Winter Sports* 2018, 2, 15.
4. Di, H.N. On the digital technology in the protection of ancient buildings. *Cult. Relics Apprais. Apprec.* 2018, 3, 130–132.
5. Koeva, M.; Luleva, M.; Maldjanski, P. Integrating Spherical Panoramas and Maps for Visualization of Cultural Heritage Objects Using Virtual Reality Technology. *Sensors* 2017, 17, 829. [CrossRef] [PubMed]
6. Kuznetsova, I.; Kuznetsova, D.; Rakova, X. The use of surface laser scanning for creation of a three-dimensional digital model of monument. *Procedia Eng.* 2015, 100, 1625–1633. [CrossRef]
7. Han, L.F.; Hou, Y.D.; Wang, Y.J.; Liu, X.B.; Han, J.; Xie, R.H.; Mu, H.W.; Fu, C.F. Measurement of velocity of sand-containing Oil-Water two-phase flow with super high water hold up in horizontal small pipe based on thermal tracers. *Flow Meas. Instrum.* 2019, 69, 101622. [CrossRef]
8. Gupta, M.; Yin, Q.; Nayar, S.K.; Iso, D. Structured Light in Sunlight. In Proceedings of the IEEE International Conference on Computer Vision, Sydney, Australia, 3 March 2014; pp. 545–552.

9. Gelabert, P. Comparison of fixed-pattern and multiple-pattern structured light imaging systems. *Proc. Spie Int. Soc. Opt. Eng.* 2014, 8979, 342–348.

10. Kinect Outdoors. Available online: www.youtube.com/watch?v=rI6CU9aRDlo (accessed on 2 December 2015).

11. O’Toole, M.; Mathet, J.; Kutulakos, K.N. 3D Shape and Indirect Appearance by Structured Light Transport. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Columbus, OH, USA, 23–28 June 2014.

12. Steger, C. An unbiased detector of curvilinear structures. *IEEE Trans. Pattern Anal. Mach. Intell.* 1998, 20, 113–125. [CrossRef]

13. Usamentiaga, R.; Molleda, J.; García, D.F. Fast and robust laser stripe extraction for 3D reconstruction in industrial environments. *Mach. Vis. Appl.* 2012, 23, 179–196. [CrossRef]

14. Haverinen, J.; Röning, J. An Obstacle Detection System Using a Light Stripe Identification Based Method. In Proceedings of the IEEE International Joint Symposia on Intelligence and Systems, Rockville, MD, USA, 23 May 1998; pp. 232–236.

15. Zhang, L.G.; Sun, J.; Yin, G. Across Structured Light Sensor and Stripe Segmentation Method for Visual Tracking of a WallClimbing Robot. *Sensors* 2015, 15, 13725–13751. [CrossRef] [PubMed]

16. Wiggins, R.A. Minimum entropy deconvolution. *Geoeplorafion* 1978, 16, 21–35. [CrossRef]

17. Mcdonald, G.L.; Zhao, Q.; Zuo, M.J. Maximum correlated kurtosis deconvolution and application on gear tooth chip fault detection. *Mech. Syst. Signal Process.* 2012, 33, 237–255. [CrossRef]

18. Mathis, H.; Douglas, S.C. Bussgang blind deconvolution for impulsive signals. *IEEE Trans. Signal Process.* 2003, 51, 1905–1915. [CrossRef]

19. Collins, R.T.; Liu, Y.; Leordeanu, M. On-line selection of discriminative tracking features. *IEEE Trans. Pattern Anal. Mach. Intell.* 2005, 27, 1631–1643. [CrossRef] [PubMed]

20. Kai, Q.; Xiuhong, C.; Baiwei, S. Robust fast tracking via spatio-temporal context learning. *Comput. Eng. Appl.* 2016, 52, 163–167.

21. Yuehua, L.; Jingbo, Z.; Fengshan, H.; Lijian, L. Sub-pixel extraction of laser stripe center using an improved gray-gravity method. *Sensors* 2017, 17, 814.

22. Sun, J.H.; Wang, H.; Liu, Z.; Zhang, G.J. Fast extraction of laser strip center in dynamic measurement of rail abrasion. *Opt. Precis. Eng.* 2011, 19, 690–695.

23. Li, Y.; Li, Y.F.; Wang, Q.L.; Xu, D.; Tan, M. Measurement and defect detection of the weld bead based on online vision inspection. *IEEE Trans. Instrum. Meas.* 2010, 59, 1841–1849.

24. Zhang, L.; Ke, W.; Ye, Q.; Jiao, J. A novel laser vision sensor for weld line detection on wall-climbing robot. *Opt. Laser Technol.* 2014, 60, 69–79. [CrossRef]

25. Furukawa, Y.; Ponce, J. Accurate, Dense, and Robust Multiview Stereopsis. *IEEE Trans. Pattern Anal. Mach. Intell.* 2010, 32, 1362–1376. [CrossRef] [PubMed]