Simultaneous shape registration and active stereo shape reconstruction using modified bundle adjustment

Ryo Furukawa
Hiroshima City University, Japan
ryo-f@hiroshima-cu.ac.jp

Genki Nagamatsu, Hiroshi Kawasaki
Kyushu University, Japan
{nagamatsu.genki.310@s,kawasaki@ait}.kyushu-u.ac.jp

Abstract

Simultaneous registration and shape fusion using 3D scanners have been proposed for conducting wide-area and dense 3D shape reconstruction. However, because the 3D scanners for such a system must be robust and should provide feedback in real time, only a few devices are available, thereby limiting the application of the technique. In this study, we propose a new wide-area scanning algorithm that only requires an off-the-shelf projector and a camera. In our technique, the devices are not necessarily fixed to each other and the relative positions of the devices as well as the scene shapes can be precisely estimated by bundle adjustment (BA) in case of structured light. To efficiently perform shape registration, a robust and dense shape reconstruction is required, which is currently considered to be an open problem for structured light systems. In this study, we suggest a novel network-based feature detection algorithm as well as shape fusion algorithm for the solution.

1. Introduction

Dense 3D shape reconstruction of large area is considered to be important for various applications such as human body capture for medical systems, indoor room modeling for augmented reality systems, and entire shape acquisition for industrial purposes. To achieve this, simultaneous registration and shape fusion of sequentially scanned 3D data have been extensively researched and developed [26, 21, 22]. For these algorithms, 3D sensors are freely moved during the scan to cover a wide-area, therefore, a real-time 3D scanning capability is required. However, it is still not an easy task and only a few commercial sensors fit the requirement [20, 1, 14], thereby, with only a small number of options available for users.

As an alternative, the development of a projector and a camera system using off-the-shelf devices can be a practical option. However, those scanning systems have not been commonly used for wide-area scanning because of several reasons. For some potential applications, one main reason is the practical difficulty that is associated with fixing the projector and camera with each other during the dynamic scanning process. For example, if a video projector is used, it is usually large and heavy and difficult to be fixed tightly under the condition of SLAM-like process, where the system is installed on a vehicle and encounters many bumps and vibrations. Another typical condition can be found in medical systems, e.g., endoscope or laparoscope systems. Those systems are too small and complicated, and thus, it is often impossible to tightly fix the pair of the devices. If the relative position between the projector and the camera is shifted, the pre-calibrated parameters will not be used, thereby resulting in the failure of the entire scan.

In this study, we propose a new multi-frame reconstruction method which does not require the relative position of the devices to be fixed, but estimates all the parameters based on the bundle adjustment (BA). Although the original BA for cameras assumes a certain amount of stable correspondences, it cannot be applied to the structured light systems because the projectors cannot be used to capture images. To solve this problem, we introduce a shape fitting algorithm that is inspired by the iterative closest point (ICP) algorithm and the ICP cost is jointly minimized through the process. After applying BA, since multiple shapes are precisely registered, they are fused into a global shape by truncated signed distance function (TSDF) in our method.

In the paper, we also propose a oneshot shape-reconstruction technique based on grid-based active stereo [15, 31] with a learning-based approach using a convolutional neural network (CNN) and an efficient shape interpolation technique using an anisotropic radial basis function (RBF). In our method, we design a single CNN, which is expected to reduce the computational cost as well as increase the accuracy. To increase the density as well as high frequency, we introduce an anisotropic radial basis function (RBF)-based shape densifying algorithm. The contributions of our study can be given as follows:

1. a new BA technique for structured light systems to re-
construct a consistent wide-area scene by performing sequential scans is proposed;

2. a new grid-based active stereo technique using multiple features, where two-directional parallel lines and the codes are simultaneously detected from the captured patterns using a single CNN, is proposed; and

3. a new shape-interpolation and integration technique using an anisotropic RBF and a weighted TSDF technique is proposed.

One practical application of our study is a 3D endoscope in which the pattern projector cannot be fixed to the head. We built a micro-sized projector and demonstrated a 3D endoscopic system to successfully recover large areas.

2. Related Work

The structured light technique has been frequently used for practical 3D shape-capturing purposes [29, 32, 23]. Temporal and spatial encoding approaches are the two primary approaches that are used to encode the positional information into patterns. Because temporal encoding requires multiple images, it is not suitable to install the system in moving devices. Conversely, spatial encoding requires only a single image, and it can be implemented in dynamic and moving devices [20, 15, 19, 31]. Therefore, by employing such techniques, a projector-camera system can be freely moved throughout the scanning process to achieve wide-area shape scan. However, it is usually challenging to fix an off-the-shelf projector to a camera. There is an interesting solution for solving the aforementioned difficulty that involves a change in the devices’ relative positions rather than fixing them [8, 7]. However, this technique assumes dense and precise correspondences between the camera and projector; therefore, temporal coding is required and cannot be applied to our purpose.

As we will explain later, our technique has resemblance to ICP algorithm, in which multiple shapes are registered by minimizing distances between closest points [2]. Although the original ICP algorithm does not estimate the scale changes, there have been several reports on techniques that consider the scale changes in the ICP framework [6, 18]; further, the scale changes and shape distortions that are caused by the calibration errors are compensated. To efficiently estimate those parameters, Garcia et al. [11] used gray-code projection for obtaining dense correspondences between multiple projectors and cameras and used BA. Our technique is inspired by their report; however, we assume dynamic motion of either a projector or a camera, and thus, their methods cannot be applied.

Another problem of spatial encoding methods is their instability because that positional information is encoded into small regions and the patterns tend to be complicated and are easily affected and degraded by the environmental conditions. To avoid such limitations, some techniques are observed to be based on geometric constraints rather than decoding [16, 27, 24, 17, 31]. However, because such techniques are still dependent on pattern detection, their results are degraded to some extent. Although there are methods that are available to compensate for such degradation of results [12, 13], they assume that multiple images are captured by projecting multiple patterns; thus, these methods cannot be applied to perform a one-shot scan. Recently, a solution was proposed for the subsurface scattering objects [10, 28]; however, it required a specifically designed pattern, whereas a general technique that used extensively varying patterns was in considerable demand. In this study, we use fully convolutional neural networks (FCNNs) that can be referred to as U-Nets [25] for decoding 2D grid-like patterns that are projected onto the target. A U-Net is an FCNN architecture that receives an image and produces a labeled image. It is known to outperform previous FCNN architectures, including the sliding window convolutional networks [4] in segmentation tasks for medical images. Song et al. [30] proposed to decode an active stereo pattern using a CNN.

In the final step, the shape integration of multiple scans is required. Further, the signed distance field (SDF) representation has been extensively used [5] and TSDF has been recently proposed to achieve real-time system [21]. However, both techniques erase high frequency shapes.

3. Overview

3.1. System configuration

Our proposed 3D measurement system comprises a projector and a camera, as depicted in Fig. 1(left). We assume that the devices are not necessarily tightly fixed to each other. The temporally or spatially encoded patterns are projected by a projector and captured by a camera; further, these patterns are decoded to retrieve the correspondences between the image and the pattern by the algorithm. In the experiments, we used grid oneshot scan, which will be explained in the next section. To show the strength of our method, we actually developed an endoscopic system (Fig. 1(right)) in our experiments. In the system, a fiber-shaped micro pattern projector (a diffractive optical element (DOE)-based pattern projector) is inserted through the instrument channel of the endoscope; the projector slightly protrudes from the endoscope head, as depicted in Fig. 2(a), and emits structured light. Because the micro pattern projector is smaller than the channel, the relative position and orientation between the projector and camera varies through scan.

3.2. Grid-based active stereo for endoscope

In the endoscopic system, we use the “grid pattern with gapped codes” described in [10], which is based on grid pat-
3.3. Basic procedure of the method

A sequence of images is captured by the camera while the structured light is projected. Further, the 3D shapes are recovered for each frame. The 3D reconstruction of each frame comprises the three stages of pattern decoding, auto-calibration, and 3D reconstruction, which are presented in the flowchart in Fig. 3. The pattern decoding stage is processed by CNNs that are trained to extract the grid-like structures and the gap codes in the captured images. The different types of features are simultaneously extracted using a single CNN. In the auto-calibration stage, the camera parameters are self-calibrated by the method proposed in [9] using the special markers detected by the CNN. This is conducted because the projector and camera are not fixed to each other. In the 3D reconstruction stage, the extracted grid structures and code information are analyzed; further, the IDs of all the detected vertical lines are decided by the grid-based reconstruction method that has been previously described, and the 3D curves are reconstructed using a light sectioning method. All the estimated parameters and recovered shapes are used as an initial value to conduct BA to refine the 3D shapes as well as the calibration parameters of the projector-camera system, that can be referred to as active bundle adjustment (active-BA). Finally, the integrated shapes are constructed by the KinectFusion-like algorithm based on TSDF.

4. Bundle Adjustment for Structured Light

4.1. Overview

A structured light system is used to capture a target scene multiple times while the projector and camera system is freely moved. Here, we describe the condition of our shape acquisition, registration and integration process.

1. The target scene is rendered by a combination of data obtained from the pattern projector and camera. Further, the correspondences between the projected pattern and the captured image can be obtained by analyzing the image, i.e., using a grid-based method [15].
The obtained correspondences is sparse because of the feature of grid-based reconstruction.

2. The relative pose between the projector and the camera can be uncalibrated.
3. The same scene is captured multiple times while assuming independent motion of the camera and the projector. We can denote a single capture of the scene as a frame. There are no specific constraints that are imposed on the motions of the camera and projector.
4. The textural information of the scene is not used, i.e., the camera is used only for capturing the projected pattern, but not the textures of the scene, because projected pattern is usually stronger than the textures.
5. In our technique, we do not implement explicit loop detection algorithm, so if a drift becomes so large that correspondences cannot be found, our technique will fail. If correspondences are found, our technique efficiently spread the accumulated errors to entire scene evenly to achieve consistent shape reconstruction.

The objective of this study is to generate a integrated shape when multiple shape measurements are being consistently merged. This should be applied, for example, on the described 3D endoscopic system or any projector-camera system that is used for measuring the 3D shape of a wide area while a projector and a camera cannot be tightly fixed each other.

A naive method to achieve the objective would be to generate 3D reconstruction frame-by-frame and to align the 3D reconstructed shapes using an ICP algorithm. If the projector-camera poses are uncalibrated, they can be auto-calibrated using the projector-camera correspondences. However, such an approach is problematic. The auto-calibration results exhibit genuine ambiguity in terms of scales. Furthermore, the low-precision relative poses between the projector and the camera cause distortion in 3D reconstruction. Because of scale inconsistencies and shape distortions, the shapes from different frames do not fit consistently even after the ICP algorithm is applied.

Instead, we use a BA technique, where the information obtained from the measurements of multiple frames is used as a single solution. Unlike the BA that is conducted when only a camera is used, the texture of the scene cannot be used in BA because the projector cannot be used to capture an image. Therefore, we cannot use the information from “a single point captured from different frames” (note that the projected patterns are not fixed to the scene). Overall, the BA technique cannot be applied without performing any modification.

By considering the defined objective and its associated problems, the only usable constraint between different frames is that their shapes should be fit by a rigid transformation. The ICP algorithm has been a common tool for dealing with this type of constraint. Therefore, we propose the usage of an iterative algorithm when the ICP algorithm and BA techniques are combined. In the proposed process, the shapes of different frames are aligned with the ICP algorithm, and inter-frame correspondences are obtained with the ICP-inspired method. Using the inter-frame correspondences, BA is processed to obtain a better fitting between the shapes of different frames. These processes are repeated until convergence of BA. We can refer to the proposed algorithm as active-BA.

### 4.2. Active bundle adjustment

The steps of active-BA can be given as follows:

**Step 1** The initial parameters of the relative pose between the projector and the camera as well as the positions of the shapes of the frames are given as input.

**Step 2** The 3D shape for each frame is reconstructed from the current pose information.

**Step 3** The corresponding points between different frames are sampled using a proximity relation between the frame surfaces, which is similar to the process that is used to retrieve the corresponding point pairs in the original ICP algorithm.

**Step 4** Using the obtained inter-frame corresponding points, the corresponding intra-frame information, which is assumed to be known by spatial coding, is propagated to another frame.

**Step 5** Using the corresponding intra- and inter-frame information, an algorithm which minimizes both of the reprojection errors of 3D points within each frame and distances between the corresponding points between different frames. Using all the output poses, dense 3D shapes and the relative position between the frames are updated.

**Step 6** The above steps are repeated until convergence.

For the relative pose of Step 1, the precision of the initial relative pose may be low; however, it should be sufficiently precise to achieve convergence.

In Step 2, if the correspondences between the projected pattern and the camera image is sparse, the reconstructed shape becomes sparse. It is difficult to align the sparse shapes because the correspondence points between

![Figure 4. Active bundle adjustment correspondence finding process. This process is equivalent to ICP.](image-url)
the frames are observed to be small (note that the ICP algorithm does not work between sparse points). Thus, we apply a shape interpolation method to densify the reconstructed shapes. In our method, normal vectors are obtained by 2D regression for each point of the coarse shape data, and RBF-based interpolation is applied using the normal information.

Step 3 is described using Fig. 4. Let the camera and the projector of frame $k$ be $C_k$ and $P_k$, respectively. By processing the image of frame $k$ (captured with $C_k$), the 2D correspondences between $C_k$ and $P_k$ are obtained. Let the $j$th pair of correspondences be a pair of $u^C_{kj}$ of $C_k$ and $u^P_{kj}$ of $P_k$. The 3D point obtained by the triangulation of $u^C_{kj}$ of $C_k$ and $u^P_{kj}$ of $P_k$ can be denoted by $p_{kj}$. Further, for all the correspondence pairs, frame $k$ is reconstructed (i.e., repeated for $j$).

Let frame $l$ be another frame; further, all the correspondence pairs of frame $l$ are also reconstructed. If the reconstructed points $p_{kj}$ are sparse, they should be interpolated, and the depth image $D_l$ with the view $C_l$ is obtained. Further, $p_{kj}$ is projected onto $D_l$ using the pose and intrinsic parameter of $C_l$. If the projected pixel is a valid 3D point, we can define this point in the corresponding point of $p_{kj}$ in frame $l$. This corresponding point is $\pi_l(p_{kj})$. The 2D projection of $\pi_l(p_{kj})$ can be calculated using the camera $C_l$ and projector $P_l$. Let these 2D points be $v^C_{kj,l}$ and $v^P_{kj,l}$, respectively. $p_{kj}$ and $\pi_l(p_{kj})$ are corresponding points between different frames. Generally, they are different but are the neighboring points of frames $k$ and $j$.

In our algorithm, we calculate BA-style reprojection errors of the points $p_{kj}$ and $\pi_l(p_{kj})$, respectively within each frames ($k$ and $l$), and the distance errors between the corresponding points. Then, the total cost to be minimized is the weighted sum of reprojection errors of all points $p_{kj}$ and distance errors of all the pairs of $p_{kj}$ and $\pi_l(p_{kj})$.

The cost function to be minimized is as follows:

$$L(I, E, P) = \sum_k \sum_j \{\text{reproj}(p_{kj}; I_{C_k}, E_{C_k}) + \text{reproj}(\pi_l(p_{kj}); I_{P_l}, E_{P_l})\} + w_c|p_{kj} - \pi_l(p_{kj})|^2 + w_b\{S(E) - \text{Const}\}^2$$

where $I_{C_k}$ and $E_{C_k}$ are intrinsic and extrinsic parameters of camera $C_k$, $I_{P_k}$ and $E_{P_k}$ are intrinsic and extrinsic parameters of projector $P_k$, $\text{reproj}(\cdot)$ is BA-style reprojection errors, $I$ is the set of intrinsic parameters $I_{C_k}$ and $I_{P_k}$, $E$ is the set of extrinsic parameters $E_{C_k}$ and $E_{P_k}$, and $P$ is the set of $p_{kj}$ and $\pi_l(p_{kj})$. $L(I, E, P)$ is minimized with respect to $I, E$ and $P$. $S(E)$ is a scale function that determines the scale of the scene, and Const is a constant value. For example, $S(E)$ can be the sum of distances between the positions of the camera or projector for all the frames. We can use the sum of distances for randomly sampled devices and frames to reduce computational costs. $w_c$ and $w_b$ are weights for cost terms. Without the term $w_b(S(E) - C)^2$ that fix the scale of the scene, the whole scene may shrink with similarity transform infinitely, since shrinking the scene reduces the distances $|p_{kj} - \pi_l(p_{kj})|$ without effecting reprojection functions.

In real applications, such as the endoscopic systems in this paper, intrinsics parameters of the camera for all frames $k$ is the same. Thus, we use a common intrinsics for all $k$.

The active-BA can be regarded as a variation of the ICP algorithm, where ICP is used to estimate only a rigid transformation between the frames; however, in this study, the proposed algorithm estimates the projector-camera relative pose, which deeply affects the shapes of the frames.

5. Implementation

5.1. Simultaneous detection of grid structures and codes

We propose the extraction of grid structure and gap-code information using U-Nets [25]. The structure of the network is depicted in Fig. 5. The numbers in the figure represent the dimensions of the feature maps. Due to this network structure, both the fine and coarse resolution features can be learned. By applying U-Net to an image, feature maps can be generated in the same size as that of the input image. In Fig. 5, the output layers are divided into three groups, each comprising the vertical, horizontal, and gap-codes.

The training process of U-Net for vertical-line detection is as follows. First, the image samples of the pattern-illuminated scene are collected. Further, the vertical line locations for the image samples are manually annotated as curves of 1-dot widths. Because the 1-dot width curves are considerably narrow to be directly used as labeled regions of teacher data, regions with 5-dot width to the left and right sides of the thin curves are extracted and are labeled as teacher data (Fig. 6). The line detection and code detection are simultaneously processed using a single U-Net. The training data for code classification are depicted in Fig. 6.
The U-Net is trained using the loss function of the softmax entropy. In this training process, we augment the training data by adding noise and scaling the intensity because the intensity of the illuminated patterns may considerably change. Examples of the U-Net output are depicted in Fig. 7. Further, the dark, bright, noisy images with different grid sizes and up to 40-degree rotation are processed accurately.

By applying the trained U-Net to the image, the 3-way labeled image for vertical line detection is obtained, where the left and right sides of the vertical curves are labeled as 1 and 2, respectively (the green and red regions in Fig. 7). By extracting the borders between these regions, curves can be detected.

The classification of gap codes can be processed by directly applying U-Net to the image signal and not from the line detection results. Thus, the gap code estimation does not depend on line detection, which is an advantage because the line detection errors do not propagate to code decoding.

5.2. RBF-based shape interpolation

To achieve stable calculation of the inter-frame correspondences, the sparse 3D points that are obtained from the sparse projector-camera correspondences should be densified. To achieve this, we use an RBF for interpolation of the 3D curves [3].

The interpolation sizes are defined by RBF functions. In case of using an isotropic RBF, to interpolate the pixels between adjacent lines of the grid, the scale parameter of the Gaussian should be larger or approximately equal to the apparent sizes of the grid in the captured images. However, this setting often oversmooths the measured shape features of the 3D curves, which can be accurately reconstructed by the light sectioning method.

To overcome this oversmoothing problem, we propose the use of an anisotropic RBF in which the kernel shape is observed to be long with respect to the direction that is perpendicular to the grid line. In a simple case in which the reconstructed lines are vertical in the image plane, we use a Gaussian kernel that is narrow in the vertical direction (anisotropic RBF). In this case, the vertical size of the kernel becomes 0.3 times of the original size. Further, the shape features along the line direction are expected to be preserved more than isotropic RBF.

6. Experiment

6.1. Evaluation of active bundle adjustment for a structured light system

To evaluate active-BA, we capture a scene multiple times (19 sets in the experiment) using the projector-camera system. A grid pattern is projected onto the object and the images are captured by the camera. The relative positions of the projector and the camera are slightly different for each frame to simulate the loose connection.

Under these conditions, auto-calibration and 3D reconstruction are processed for each frame. Because of frame-wise auto-calibration, the scales of different frames are inconsistent; further, calibration errors are also present. To create meaningful results from the auto-calibration method, we set a constant value as the baseline for all the frames. To emphasize the calibration errors, we also added Gaussian noise to the extrinsic parameters.

Using these data, the result of ICP alignment for the initial shapes and the result of the active-BA are shown in Fig. 8. While the result of aligning the initial shape exhibits considerable differences between different frames, the differences that are present in the result of active-BA are drastically minimized. We also scan the same object using KinectFusion [21] and results are shown in Fig. 10. Since there is no loop closure mechanism on KinectFusion, inconsistent shapes are reconstructed after entire shape scan.

For performing the evaluation, we apply plane fitting to...
Figure 8. Active bundle adjustment results: (left to right) two captured images, before and after active bundle adjustment.

Figure 9. (Left) Camera and projectors’ positions and reconstructed points before/after BA. Yellow/cyan for cameras, orange/azure for projectors and red/blue for points. (Right top) 3D points before BA and (right bottom) after BA.

Figure 10. (a) single captured shape by Kinect v1, (b) KinectFusion result with single rotation around the object, (c) two times rotation and (d) three times rotation. Note that since there is no explicit loop closure mechanism on KinectFusion, a large inconsistency in shape remains even if several rotations are conducted.

Figure 11. Grid and code detection results for seashell-shaped soap: (a) the appearance of the sample; (b) the captured image; and (c) extracted grid structures and codes using U-net.

Figure 12. Comparison with the ground truth of a seashell-shaped piece of soap: (a) ground truth captured by gray code; (b) our method with anisotropic RBF; (c) our method with isotropic RBF.

Table 2. RMSE of each result.

<table>
<thead>
<tr>
<th></th>
<th>Fig. 12(b) Anisotropic RBF</th>
<th>Fig. 12(c) Isotropic RBF</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSE</td>
<td>1.60mm</td>
<td>1.74mm</td>
</tr>
</tbody>
</table>

6.2. Evaluation of the effect of anisotropic RBF and TSDF-based shape fusion

To evaluate the anisotropic RBF, we scan a seashell-shaped piece of soap (Fig. 11) including isotropic RBF, and compare the results with the ground truth. The results are presented in Fig. 12 and Table 2. Our proposed method that uses the anisotropic RBF features the finest details, and its RMSE is the optimal as compared to isotropic RBF.

6.3. Active BA and shape integration with TSDF using the endoscopic system

We apply our online shape registration and merging algorithm to the static objects, i.e., phantom of stomach, and results are presented in Fig. 13. In Fig. 13(top), the area of the recovered shape from a frame of the captured sequence is shown by the blue polygon. Fig. 13(bottom left) depicts the captured image of the region highlighted by the yellow rectangle in top figure, where the grid pattern is projected to the surface. In the image, the grid lines are disconnected by the high-frequency shape of the model surface and shapes are distorted, as depicted in Fig. 13(bottom right). The integrated shape generated by active-BA and the online registration and merging algorithm is depicted in Fig. 14(bottom row). A large area is successfully recovered, and the high-frequency shape details are retained. For performing comparisons, we scan the same object with KinectFusion [21] and the result is presented in Fig. 14(top row); the shape is heavily distorted because of sparse reconstruction without consistent shape recovery algorithm. RMSE is 3.63mm for our active-BA, whereas 4.55mm for KinectFusion.

Please refer to the supplementary materials for more visual results.

6.4. Active BA and shape integration for wide area using projector and camera system

To confirm the strength of our method for common projector and camera based system, we scan the entire room using the pro-cam system as shown in Fig. 15. We rigidly...
attached projector and a camera using professional camera tools, however, since projector is heavy, there is a possibility that slight motion occurs during moving the system. More importantly, calibration of such in-out scenario is usually not easy, because it is usually difficult to prepare large calibration objects. We apply common calibration technique using natural feature points, which are extracted from the scene for initial shape reconstruction. For comparison, we conducted a projector and a camera calibration technique using BA [7]. Since the system is tightly fixed, results are mostly comparable, thereby effectiveness of our method is confirmed. Note that if the projector and the camera’s relative position is moved, the technique [7] cannot be applied.

Please refer to the supplementary materials for more visual results

7. Conclusion

We propose and demonstrate a wide-area scanning technique using an off-the-shelf projector and camera. The projector and camera are not required to be fixed with each other and can be freely moved to scan the entire shape of the target object or wide region of the scene. The shapes and relative position between the projector and the camera can be precisely estimated using our technique that can be referred to as active-BA. Because the technique uses ICP to retrieve the correspondeces between frames, a robust and dense reconstruction is required. For performing robust shape reconstruction, a single U-Net based network is simultaneously used for both line detection and code-based segmentation. Finally, using our method, wide-area 3D shapes can be successfully reconstructed from both small and large scene. In future, we intend to develop a real-time system for autonomous robot-based navigation.

Acknowledgement

This work is supported by JSPS/KAKENHI 16H02849, 16KK0151, 18H04119, 18K19824, and MSRA CORE14.
References


[14] Intel. Intel realsense SR300. 1


