Real Time Detection of Repeated Structures in Point Clouds of Urban Scenes

Sam Friedman

Hunter College & Graduate Center
695 Park Ave, New York NY, 10021
http://www.welike2draw.com/vision/

Ioannis Stamos

Hunter College & Graduate Center
695 Park Ave, New York NY, 10021
http://www.cs.hunter.cuny.edu/~ioannis

Abstract—**Laser range scanners provide rich 3D representations of urban scenes. These scenes are replete with repetitive architectural features such as windows, balconies, and cornices. Processing of dense 3D images is often computationally intensive and occurs offline after acquisition. Here, however, we present an online algorithm for the detection of repetitive features in 3D range scans. Our algorithm creates a function from each scanline by deriving a local measure at each point. Computing the Fourier transform of that function reveals the periodicity of the scene. This robust algorithm forms the basis for novel methods of feature extraction, compression, and registration. What is more this whole process can be executed on-the-fly and integrated into hardware transforming laser scanners into architecture aware devices.**

Keywords-**window detection; compression; range images; urban scenes; fourier transforms**

I. INTRODUCTION

Modeling 3D urban scenes from range scan data is a field of active research [1], [2], [3]. One of the major problems in scanning outdoor scenes is the density and complexity of the acquired datasets. Detecting prominent structures such as windows, balconies, or cornices is a required step for scene understanding. These structures can be used for compressed scene representations, registration applications and higher-level recognition tasks. The detection of these structures is currently a slow process that is performed offline after the acquistion of the whole dataset. Detection must overcome missing data and occlusions in these varied structures. However, the periodicity of these structures within the urban environment is eminent. In this paper we present novel online algorithms for the efficient extraction of repetitive features in large-scale urban scenes by exploiting their periodicity. We present a number of applications, including compression of the large-scale datasets.

The laser sensor measures distance by shooting a laser beam at an object and recording the time it takes the laser beam to return. If the laser beam does not return the scanner cannot make a measurement. For this reason windows present a challenge. Unobstructed windows often absorb the laser beam into an interior from which they do not escape. In those cases a facade is dotted with missing data. Curtains and blinds can also cause problems depending on their material. They may deflect the laser beam away

from the scanner or because of ruffles and gaps provide inconsistent measurements. While windows are unpredictable, in urban environments they tend to occur regularly. Despite the difficulty of interpreting data from a single window, a vertical line of windows will be predictably unpredictable. The areas of missing and inconsistent data will occur regularly, with a defined period. Balconies have highly variable appearances due to self-occlusions and vegetation and glass doors. However their periodicity is also recognizable.

Fourier analysis gives insight into periodic elements. Constructing the Fourier transform of a signal from a vertical line containing windows or balconies identifies the frequency and size of those features. Adjacent scanlines are likely to pass over the same features and can be used to reinforce the signal. The entire process can be executed while scanning is underway allowing for a real-time analysis of the scene. After the detection of the locations of periodic elements further processing can be performed. We start by presenting related work and then present the details of our approach.

II. RELATED WORK

Detecting repeated structures such as windows and balconies in urban scenes is a problem that has received significant attention. Using 2D images as input, shape grammars for architectural design have been used in modeling facades acquired through aerial photographs [4] and for the generation of synthetic cities [5].

A similar approach that uses ground-based images was presented in [6]. An earlier approach is the one of [7]. The work of [8] uses a sequence of images to produce a sparse 3D point cloud as input to the detection algorithms.

Fewer approaches exist when the input is a laser range scan. [9] derives regularities of substructures from a 3D model or range scan of a scene. This works by detecting symmetries (or similarity transformations) of basic structures in a regular grid. This general approach can be used for extracting regularities but it is sensitive in the calculation of curvatures and computationally intensive. In [10] window-like rectangular features were extracted by using 3D edge detection on high-resolution 3D data. Line features are again utilized in [11] for symmetry detection. [12] presents an interactive interface which exploits regular structures in urban scenes to improve sparse point clouds. Employing a

Figure 1. A texture mapped range scan. The balconies and windows seen here can be detected and extracted on-the-fly using the techniques discussed in this paper.

Markov Network approach [13] labels points as windows, but requires training.

Our *contributions* with respect to earlier work can be summarized as follows:
(a) we efficiently extract repeated structures in range datasets of large-scale scenes in an online fashion by processing the data as each scanline is being acquired, (b) by using periodicity as the basis of our detection, we can have robust results even in low resolution areas of the scene, (c) we apply our feature detection results for compression, and (d) we do not require training.

III. DATA ACQUISITION

Our lab operates a ScanStation2 3D range scanner [14]. We collect 2.5D range data of large scale urban scenes. Scans contain up to 10 million points depending on the field of view and resolution. The scans have a resolution of about 10 points per square meter at a 50 meter radius. Resolution varies with the distance from the scanner, (a detail which our algorithm must address in order to properly recognize regular structures). The scans provide a 360 degree view within a 300 meter radius encompassing about one city block, for example see Figure 1.

The scanner is placed on a steady platform and measures the distance to the closest surface. The measurement is acquired by emitting a laser beam and precisely timing the delay between when it is emitted and when it is reflected back. The incremental movement of the beam is a parameter set by the user. No distance is measured when the laser hits transparent or highly specular surfaces, or when the closest surface is at a distance greater than a threshold (300 meters in our setting). For example when the laser is pointed at the sky nothing reflects the beam and no data is recorded.

IV. THE ALGORITHM

Each column from the 2D array of data corresponds to a vertical set of 3D points. For this reason the data retrieved from the scanner is often referred to as 2.5D. We will use the terms scanline and column interchangeably as both refer to a vertical set of 3D points forming a column within the 2D array. We create a signal from this set of points. When the laser has passed over periodic features this repetition appears in the signal. By performing Fourier analysis we can extract the period of these elements. We group adjacent columns covering the same features. Region growing extracts the precise features. The range scan can be compressed with knowledge of the periodicity of a feature.

A. Preprocessing

Data points from a scanline of an urban scene typically occur near one of two major planes, the ground plane and the facade plane. The ground plane points include pavement, vegetation, people, cars and other accoutrements of a city street. These points rarely include periodic elements and are not of interest here. On the other hand, points that make up the facade plane often incorporate periodic elements and are very much of interest. Often there is another group of points of unresolved data that come from the scanline passing over the building and sending the laser into the sky from which it does not return.

Clearly we would like to separate the points that might contain periodic elements from those we are certain do not. Using [15] we can classify points which belong to the ground plane and remove them from our consideration. The unresolved data points from the sky region are also removed.

It is often the case that a scanline will cover one facade of interest and then upon reaching the top of that facade record fragments of other buildings further in the distance or equipment on the roof like water towers or central air-conditioning units. These data are not of interest when searching for the periodic elements in the facade. We ignore them by introducing a jump threshold which terminates calculation of the column function when adjacent points in the column are more than a certain distance apart.

B. The Column Function

We must devise a signal from the scanline on which we will perform the Fourier analysis. This function must reflect the regularity of the features, and therefore yield a periodic function for each column. We will compute the Fourier Transform of each of these functions to discover the regularity. See Figures 2 and 3.

The measure of the angles between neighboring points in a column is one choice. This can be computed efficiently. However since it only takes into account three points it is a noisy metric. Nevertheless the regularity of windows shines through and even with this primitive measure regularity can be deduced.
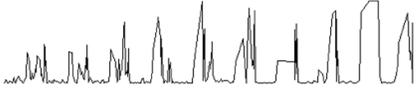
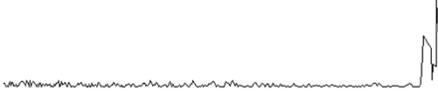Figure 2. A column function covering windows.

Figure 3. A column function that does not pass over windows.

To obtain the column function from angles we define

$X_{i,k}$: A three dimensional point located in the ith column and jth row of the scan.

$D_{i,k}$: $X_{i,k+1} - X_{i,k}$ (the vector between two successive measurements in a given scanline i).

$V_{i,k}$: the angle of the vector $D_{i,k}$ with the vector $D_{i,k+1}$.

As we obtain sequential observations along any column i, we can form the function from the sequence of angles $\{V_{ik}\}$, $k = 1, 2, \ldots$. When the sequence passes over planar areas like facades of buildings we expect the $V_{i,k}$ to be near 0, see Figure 3. As the sequence moves over non-planar and discontinuous regions such as vegetation or architectural features we expect $V_{i,k} >> 0$, see Figure 2. Using this sequence of angles [15] demonstrates a way to classify points as belonging to the horizontal ground plane, a vertical facade, or vegetation.

A more robust measure can be achieved by estimating the curvature at each point in the column. Cazals and Pouget have developed a method of estimating curvature in 3D data [16]. Using osculating polynomial jets they have devised a method to approximate the Gaussian and mean curvature. This method takes into account the local neighborhood of a point and is less subject to noise than the angle metric but more computationally expensive.

We have had success with the quicker metric of successive angles and have not resorted to the more laborious calculation of curvatures in much of this work. However, the feature extraction (V-A) algorithms relies on curvature estimation. It is also likely that in certain situations this more robust metric will prove useful.

Another potential signal can come from the Principal Component Analysis. A rough estimate of the planarity of a neighborhood is given by the eigenvalue decomposition of the covariance matrix. The smallest eigenvalue of the covariance matrix gives a measure of the planarity of the region. While facades are planar most architectural features of interest are not. The sequence of eigenvalues therefore
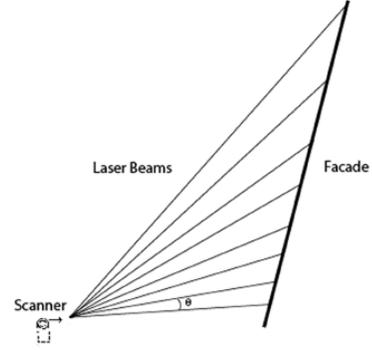

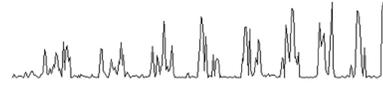
Figure 4. The uneven sampling of the range scan.

Figure 5. The distortion of the column function due to uneven sampling.

can also be used to construct the column function.

C. Interpolation

The laser scanner rotates a fixed angle between readings. However the objects measured are not equidistant from the scanner. Just as our vision skews the relative size of near and far objects, the scanner samples nearby objects with higher resolution than distant ones. The column function will reflect this fact. The periodicity of a column will be distorted by the uneven sampling of the laser. See Figure 4 and 5.

To address this fact interpolation must be performed on the column function. We use a nearest neighbor interpolation which adjusts the column function to reflect the true distances. When we compute the angle, curvature or eigenvalue that compose the column function the distance of each point to the ground plane is also computed. The points are sorted based on these distances. The column function is then resampled with a consistent distance interval. This corrects the distortion from the scanner.

D. The Fourier Transform

The Fourier transform decomposes a function into sinusoids [17]. The magnitude of the Fourier transform indicates the amount of each frequency that is present in the function. The frequency of a periodic function will register as a local maximum in the magnitude of its Fourier transform.

With this theory in mind we take the Fourier transform of the column functions. When the column has passed over a periodic element such as a series of windows this repetition appears as a local maximum in the magnitude of the Fourier transform. By extracting the local maxima in the Fourier
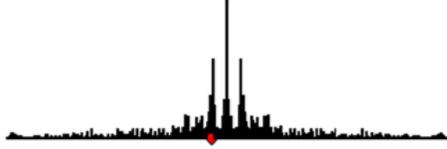
Figure 6. The Fourier Transform of a column function with windows, after interpolation. The red dot is placed under the dominant frequency. The absolute maximum is the zeroth frequency or the mean of the column function.
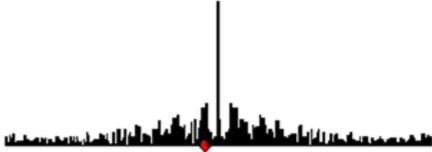

Figure 7. The Fourier Transform of a column function with windows before interpolation. Again the red dot is under the dominant frequency, but without interpolation the local maxima are more diffuse.

Figure 8. The dominant period plotted in several range scans. The red circles indicate the periods.

transform we can discover which frequencies occur in the column. We ignore the zeroth frequency which is often the global maximum because it represents the mean of the function and does not correspond to a period.

As desired the transforms of columns containing repeated windows show noticeable maxima at frequencies corresponding to the windows in the column. Figure 6 shows a typical local maxima found in periodic column functions. Conversely the transforms of windowless columns show only the global maximum of the zeroth frequency and no noticeable local maxima.

Without the interpolation which corrects the uneven sampling of the scanner, the local maxima will be more diffuse. This makes it more difficult to pinpoint the exact frequencies present in the column. Compare the transform in Figure 6 to Figure 7. Interpolation is therefore essential in obtaining the precise frequency of the scanline.

The problem of finding repeated structures has been reduced to recognizing local maxima in a 1D Fourier Transform. With one pass over the transform we can extract the relevant frequencies. The data is then segmented into pieces corresponding to the period and the repetition can be seen in the scan. See Figure 8.

E. Column Classification

As the scanner moves around the scene many columns of data are recorded. For each column we compute the Fourier transform. By subtracting the zeroth frequency from the transform we obtain a measure of the quantity of periodic elements within the scan. This measure can be used as a threshold to classify individual columns. Columns with a high measure are likely to contain periodic elements while columns with a low measure are not.
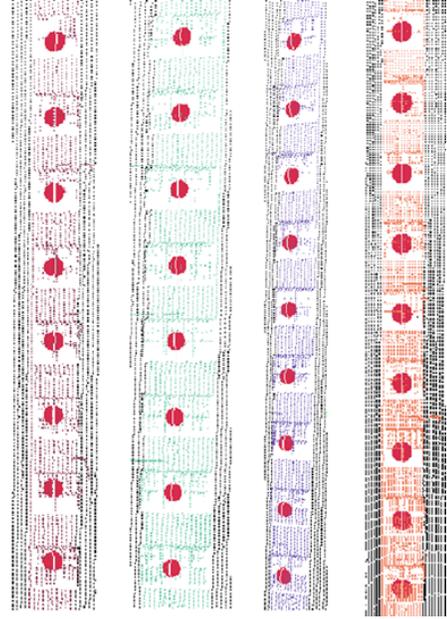
Using this quantity we can group adjacent columns based on periodicity measures. Each group of columns will correspond to a repeated feature and should encompass all the relevant data points of that feature. See Figure 9.

The column groups can also be used to buffer against noise in the individual column functions. The Fourier transforms of the column functions from the group can be summed together to make a more robust signal. However since different columns may be indexed differently the Fourier transforms must be aligned before they are summed. This alignment can be forced by centering the transforms along all of their zeroth frequencies and then summing. This aggregate transform is a smoother more robust signal than the transforms of the individual column functions. From the aggregate transform we can make a confident estimate of the period of each group of columns. See Figure 10.

V. APPLICATIONS

The rapid detection of repeated structures opens the door to an arsenal of potential algorithms. The knowledge of a periodic element and its extracted point cloud can be used for compression, registration, segmentation, and matching.

A. Feature Extraction

Once a repeated feature has been recognized, we may desire deeper insight into that feature. We know from the local maxima in the Fourier transform of the aggregated columns what periods are present within a group of columns. Now we wish to discover exactly what is repeating.
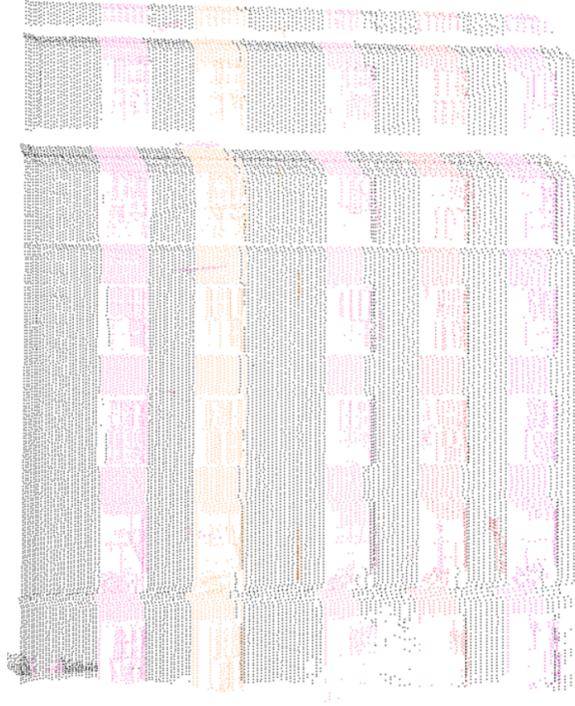
Figure 9. The columns are classified as windowed and windowless. Adjacent windowed columns are grouped. See the pdf for color.
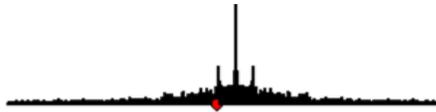

Figure 10. The Aggregated Fourier Transform of a group of adjacent column functions. The highest magnitude non-zero frequency is extracted as the dominant frequency. There is a red dot indicating the dominant frequency.

To extract the repeated feature we grow regions of points based on curvature. If the column function was constructed from curvature then the curvatures at each point are already known. Otherwise they must be computed now using the method described in [16].

Once the curvatures are known the points are grouped through a connected components algorithm. Points with high curvatures are grouped with neighboring points that also have high curvature. Planar points are ignored. The process is optimized using the disjoint sets data structure. See Figure 11.

B. Shape Signatures and Verification

To compactly identify the features we compute shape signatures for each extracted region. These shape signatures take the form of histograms as described in [18]. To generate the histogram we randomly select 5000 pairs of points within the region. For each pair we calculate the euclidean distance
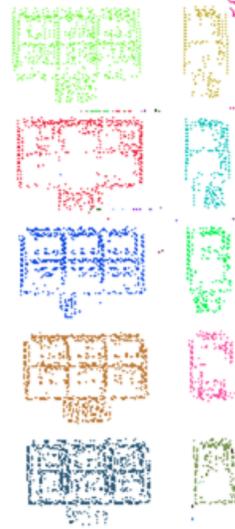

Figure 11. Some extracted features. Each distinct feature is colored differently, see the pdf for colors.

between them and create the histogram by quantizing these distances. These shape signatures can be shown to be translation, rotation, and reflection invariant.

Architectural periodicity is often the result of the repetition of a single building element, like a window or a balcony. However, it is also possible that distinct features will be periodically aligned, for example windows directly below balconies. Histograms, like those in Figure 12, allow us to verify which features are involved in the repetition the transform exposed. Similar features will have similar histogram signatures. Comparing 3D shapes is a challenging problem. Comparing histograms, however, is easy, we simply sum their differences. The calculation is quick and gives a decent estimation of the similarity between the two shapes.

C. Compression

The knowledge of an exact repeated feature lights the way for compression. Fourier analysis indicates if a group of columns is periodic. This knowledge alone is sufficient for a primitive and lossy compression. The group is segmented into regions that have a height equal to the distance of one period as determined by the Fourier transform. Because windows typically have vertical and horizontal symmetry we wish to extract the feauture's surrounding square. The Chebyshev distance between two points α and β in $\mathbb{R}^3$ is defined as

$$D_{Chebyshev}(\alpha, \beta) : max_{i \in (1,2,3)}(|\alpha_i - \beta_i|).$$

We use this metric to determine the square region with sides equal to the length of the dominant period. One of these regions is chosen as the representative, in our experiments we choose the densest feature from the set in V-A. To
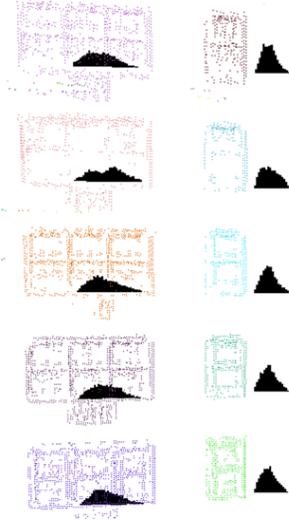
Figure 12. The histograms computed for each features. Notice how similar features have similar histograms.
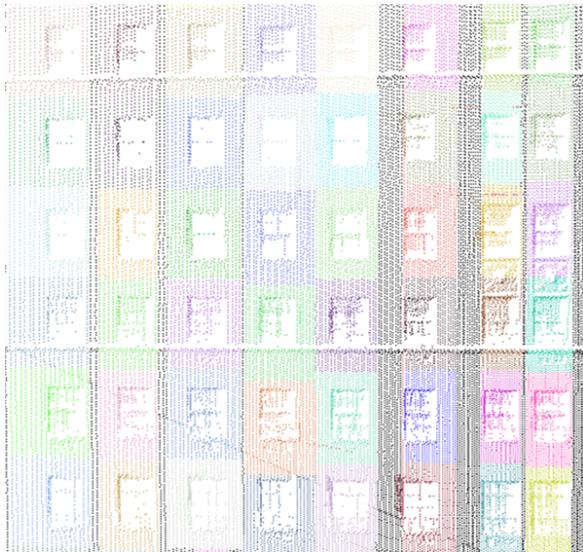

Figure 13. The columns containing windows are segmented by their dominant period. See the pdf for color.

complete the compression we translate the chosen region vertically by the distance of the period. In this way the size of the column can be reduced by the number of windows found in the group of columns.

For a more ambitious reduction we can assume that the non-repeating areas of the building are planar. This is often, but not always, true for urban structures. The vertical distance that came from the Fourier analysis can be coupled with the horizontal distance to the next adjacent group of periodic columns. The Chebyshev distance metric is altered to handle two different distances one vertical, the other

horizontal. The scanner determines the third basis of the coordinate system to be in the vertical direction. Therefore the vertical axis corresponds to the third coordinate of the points. The horizontal direction, however, is not necessarily a basis so we take the Chebyshev distance of both the first and second coordinates.

$D_{Vertical}(\alpha, \beta)$: $(|\alpha_3 - \beta_3|)$
$D_{Horizontal}(\alpha, \beta)$: $max(|\alpha_1 - \beta_1|, |\alpha_2 - \beta_2|)$

Now much larger swaths of the building are included in the compression. The entire facade is compressed by a factor equal to the number of periodic elements it contains.

This naïve compression may achieve order of magnitude reductions but not without a cost. It is common for buildings to contain one type of repeated feature vertically aligned with an entirely different feature. For example windows of various sizes or windows and balconies in the same column. To account for this the compression described above is augmented to analyze what type of feature is being used for compression and to ensure that this feature is only used to replace similar features.

To achieve this the histogram shape signatures developed in V-B are computed. Now instead of blindly replacing one region with another we first compare the histograms associated with each region. If the shape signatures are similar we can continue with the compression. However if the histogram differences are above a threshold we retain both regions in the compressed scan. This way we avoid mistakenly compressing two different features into one, solely because the features were vertically aligned and exhibited some periodicity.

This type of compression is more computationally intensive as the histograms of each region must be computed. Depending on the degree of compression desired and the style of buildings in the scan users may choose to eschew this step for speed or incorporate it for enhanced reliability.

The scan depicted in Figure 14 contains 59,749 data points. Using this compression method we were able to remove 41,660 points or about 70% of the data.

D. Registration

Knowledge of the periodicity of scans can assist in registering multiple scans in the same coordinate system. Many registration algorithms attempt to register the raw data in the scans which can be computationally expensive and require user input.

Rather than dealing directly with the data, one can use the periods discovered by the Fourier transformations. We reduce the point cloud to a grid of repeated structures. For example, see the grid in Figure 15. Each column in the grid has its associated period and feature. The grid of repeated features is then registered with another scan's grid of features.

For verification the histograms associated with each architectural element are also compared. From this registration
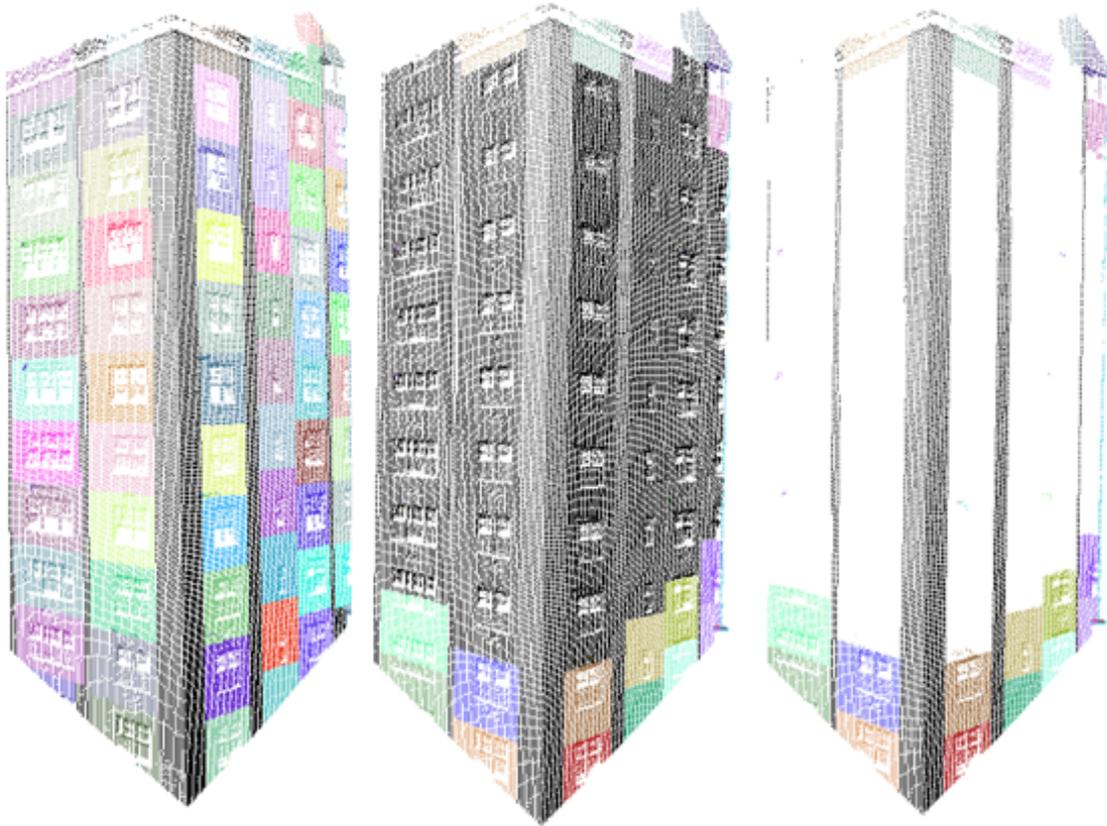
Figure 14. On the left the columns containing windows and the surrounding planar regions are segmented by their dominant period. In the center the compressed version of the scan created by translating the representative repetitive feature. On the right all but one period have been removed, showing the amount of reduction. The compressed version of the scan seen above is 70% smaller than the original.

the transformation matrix can be computed and applied to the data points to register the scans. This entire process can happen during acquisition perhaps opening the door for simultaneous recording and registration of 3D images from multiple scanners.

VI. CONCLUSION AND FUTURE WORK

These algorithms were executed on a plethora of scans retrieved by the ScanStation2 laser scanner. The results show success in recognizing the periods of repeated features, extracting those features, and performing compression.

The successive angles between points can be computed in constant time. Extracting the dominant frequency in a transform takes only one pass over the transform or $O(n)$ where n is the size of a scanline. So the complexity of the algorithm presented here is $O(n \log_2 n)$ from computing the Fourier transform with a FFT implementation.

The compression results could be improved by using the Iterative Closest Points algorithm. A small error in the detected period will be multiplied every time the representative feature is translated by that period. In large buildings this

could become noticeable. To mitigate this, every translation could use ICP to register the repeated feature against the original.

We presented an online algorithm for the detection of repetitive features in 3D range data. The application of Fourier analysis to recognizing repetitive features allows on-the-fly classification, compression, and understanding of periodicity in range images. It is our hope that algorithms of this type will continue to be developed and integrated into hardware adding an appreciation of architectural nuance to the laser detection of urban scenes.

REFERENCES

[1] P. K. Allen, I. Stamos, A. Troccoli, B. Smith, M. Leordeanu, and S. Murray, "New methods for digital modeling of historic sites," *IEEE Computer Graphics and Applications*, vol. 23, no. 6, pp. 32–41, 2003.
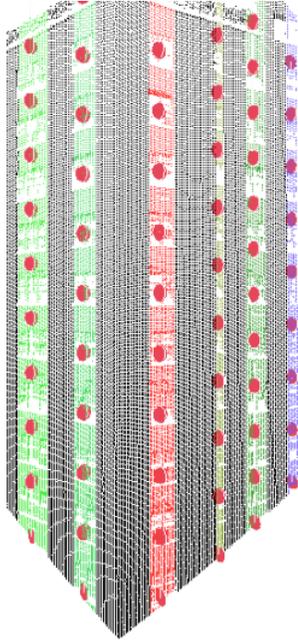
Figure 15. The range scan can be reduced to a set of repeated features. The grid of repeated features is overlaid on top of the scan. Registering regular grids is far easier than registering complex point clouds.
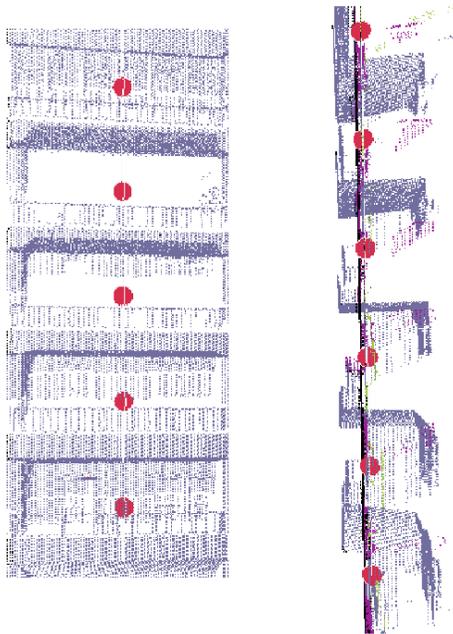

Figure 16. The algorithm can detect balconies as well as windows. On the left the balconies are seen straight on with their period depicted by the red circles. On the right the same balconies are shown in profile view. Balconies are notoriously difficult to identify because of self occlusions and their varied contents. This algorithm finds the regularity. These balconies can be seen in the center of Figure 1.

[2] I. Stamos, L. Liu, C. Chao, G. Wolberg, G. Yu, and S. Zokai, "Integrating automated range registration with multiview geometry for the photorealistic modeling of large-scale scenes," *International Journal of Computer Vision*, vol. 78, no. 2–3, pp. 237–260, July 2008.

[3] H. Zhao and R. Shibasaki, "Reconstructing a textured CAD model of an urban environment using vehicle-borne laser range scanners and line cameras," *Machine Vision and Applications*, vol. 14, no. 1, pp. 35–41, 2003.

[4] P. Muller, , G. Zeng, P. Wonka, and L. V. Gool, "Image-based procedural modeling of facades," in *SIGGRAPH*, 2007.

[5] P. Muller, P. Wonka, S. Haegler, A. Ulmer, and L. V. Gool, "Procedural modeling of buildings," in *SIGGRAPH*, 2006.

[6] O. Teboul, L. Simon, P. Koutsourakis, and N. Paragios, "Segmentation of building facades using procedural shape priors," in *CVPR*, Jun. 2010, pp. 3105 –3112.

[7] S. C. Lee and R. Nevatia, "Extraction and integration of window in a 3D building model from ground view images," in *CVPR*, vol. 2, Jun. 2004, pp. 113–120.

[8] H. Mayer and S. Reznik, "Building facade interpretation from uncalibrated wide-baseline image sequences," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 61, no. 6, pp. 371 – 380, 2007.

[9] M. Pauly, N. J. Mitra, J. Wallner, H. Pottmann, and L. Guibas, "Discovering structural regularity in 3D geometry," *SIGGRAPH*, vol. 27, no. 3, pp. #43, 1–11, 2008.

[10] I. Stamos and P. K. Allen, "Geometry and texture recovery of scenes of large scale," *Journal of Computer Vision and Image Understanding*, vol. 88, no. 2, pp. 94–118, 2002.

[11] M. Bokeloh, A. Berner, M. H.-P. Seidel, and A. Schilling, "Symmetry detection using line features," *Computer Graphics Forum, Proc. Eurographics*, 2009.

[12] L. Nan, A. Sharf, H. Zhang, D. Cohen-Or, and B. Chen, "Smartboxes for interactive urban reconstruction," *ACM Transactions on Graphics*, vol. 29, no. 4, 2010.

[13] R. Triebel, K. Kersting, and W. Burgard, "Robust 3D scan point classification using associative Markov networks," in *IEEE International Conference on Robotics and Automation*, May 2006, pp. 2603 –2608.

[14] "Leica Geosystems: Scan Station 2," 2007, http://www.leica-geosystems.com/hds/en/lgs_62189.htm.

[15] O. Hadjiliadis and I. Stamos, "Sequential classification in point clouds of urban scenes," in *3DPVT*, Paris, France, 2010.

[16] F. Cazals and M. Pouget., "Estimating differential quantities using polynomial fitting of osculating jets," *Computer Aided Geometric Design*, vol. 22, pp. 121–146, 2005.

[17] B. K. P. Horn, *Robot Vision*. MIT Press, 1986.

[18] R. Osada, T. Funkhouser, B. Chazelle, and D. Dobkin, "Shape distributions," *ACM Transactions on Graphics*, vol. 21, no. 4, 2002.