


## Article

# Learning-Based Sub-Pixel Change Detection Using Coarse Resolution Satellite Imagery

Yong Xu <sup>1</sup> , Lin Lin <sup>2</sup> and Deyu Meng <sup>2,\*</sup><sup>1</sup> Institute of Future Cities, The Chinese University of Hong Kong, Hong Kong, China; xuyong@cuhk.edu.hk<sup>2</sup> School of Mathematics and Statistics and Ministry of Education Key Lab of Intelligent Networks and Network Security, Xi'an Jiaotong University, Xi'an 710000, China; linl.xjtu@hotmail.com

\* Correspondence: dymeng@xjtu.edu.cn; Tel.: +86-181-0924-1418

Academic Editors: Qi Wang, Nicolas H. Younan, Carlos López-Martínez and Prasad S. Thenkabail

Received: 11 May 2017; Accepted: 5 July 2017; Published: 10 July 2017

**Abstract:** Moderate Resolution Imaging Spectroradiometer (MODIS) data are effective and efficient for monitoring urban dynamics such as urban cover change and thermal anomalies, but the spatial resolution provided by MODIS data is 500 m (for most of its shorter spectral bands), which results in difficulty in detecting subtle spatial variations within a coarse pixel—especially for a fast-growing city. Given that the historical land use/cover products and satellite data at finer resolution are valuable to reflect the urban dynamics with more spatial details, finer spatial resolution images, as well as land cover products at previous times, are exploited in this study to improve the change detection capability of coarse resolution satellite data. The proposed approach involves two main steps. First, pairs of coarse and finer resolution satellite data at previous times are learned and then applied to generate synthetic satellite data with finer spatial resolution from coarse resolution satellite data. Second, a land cover map was produced at a finer spatial resolution and adjusted with the obtained synthetic satellite data and prior land cover maps. The approach was tested for generating finer resolution synthetic Landsat images using MODIS data from the Guangzhou study area. The finer resolution Landsat-like data were then applied to detect land cover changes with more spatial details. Test results show that the change detection accuracy using the proposed approach with the synthetic Landsat data is much better than the results using the original MODIS data or conventional spatial and temporal fusion-based approaches. The proposed approach is beneficial for detecting subtle urban land cover changes with more spatial details when multitemporal coarse satellite data are available.

**Keywords:** land cover change; downscaling; sub-pixel change detection; machine learning; MODIS; Landsat

## 1. Introduction

Timely and accurate information about land cover dynamics is highly important for sustainable urban development and better quality of life in cities. Compared with conventional data collection methods like field surveying and aerial photography, satellite images have proven to be more effective and efficient for land use/cover change monitoring at regional or global scales due to their timely, consistent, repeatable, and cost-effective measurements [1,2]. Until now, a wide variety of change detection approaches have been formulated, ranging from preclassification methods such as image differencing, image ratioing [3], band analysis [4], principal component analysis [5], change vector analysis [6], and composite analysis to postclassification comparisons [7].

The availability of satellite data with improved spatial and temporal resolutions makes it possible to characterize land cover changes (LCCs) at higher spatial and temporal scales [8]. Some multitemporal coarse resolution (CR) sensors, such as the Moderate Resolution Imaging

Spectroradiometer (MODIS), Advanced Very High Resolution Radiometer (AVHRR), the Medium Resolution Imaging Spectrometer (MERIS), and SPOT-Vegetation, have been proven to be suitable for land use/LCC and vegetation dynamics' monitoring [9–11], with which the status and trend of land cover transitions or vegetation dynamics are characterized. Consequently, a variety of multi-temporal change detection approaches have been proposed [12–14].

CR data are effective for phenological change detection due to their high revisit frequencies, but their low spatial resolutions limit their applications for accurate monitoring of urban growth dynamics—especially for rapidly growing areas [7], where dynamic changes commonly occur in sub-pixel scales (like fields, water areas, roads). To enhance the capability of remote sensing for monitoring these dynamics at a sub-pixel scale, researchers have attempted to apply some unmixing approaches to recover high spatial resolution (HR) data directly from CR data [7,8,14–17]. In particular, Le Hégarat-Masclé et al. [8] proposed a statistically-based change detection model in which sub-pixel LCCs are estimated by utilizing previous land cover information as a reminder. Ling et al. [15,16] presented an improved sub-pixel mapping algorithm for change detection using prior land cover percentages, with which temporal contextual information was used to conduct sub-pixel change mapping. However, high-quality land cover percentages are required as input for this approach, which limits its real value. Zurita-Milla et al. [17] presented an unmixing-based approach to downscale multitemporal MERIS data for vegetation dynamics, but it is inappropriate for land cover-type changes. Though soft classification approaches can estimate land cover proportions within a coarse pixel [18], they fail to determine the spatial distribution of each class [19], and needless to say the detection of sub-pixel changes.

Another possible solution for sub-pixel change detection is to explore data-fusion approaches to obtain synthetic data with high spatial and temporal resolutions. These high-resolution synthetic data generated are then used for LCC detection at a HR. In view of conventional data fusion approaches (e.g., pan-sharpening, which integrates both spectral and spatial information rather than spatial and temporal information), they are beneficial for improving spatial resolution, but not suitable for fast change detection. Gao et al. [20] started a pioneering work to develop a spatial and temporal adaptive reflectance data fusion model (STARFM) to obtain high-quality Landsat-like data, but the underlying assumption of having no LCCs over time heavily limits its applications for seasonal change monitoring of vegetation [11,21,22]. In the meantime, Hilker et al. [21,22] proposed an improved spatial and temporal data fusion approach named STAARCH, in which an optimal Landsat was selected with a defined forest disturbance index. It is efficient for detecting forest disturbance, but not very efficient for complex LCCs in cities. Similarly, Zhu et al. [23] proposed an enhanced STARFM to extend the applications of the original approach for complex areas with heterogeneous landscapes. Roy et al. [24] presented a semiphysical fusion approach to characterize surface reflectance variation with the BRDF spectral model parameters and the sun-sensor geometry over time, but it is still not efficient for the fusion task with land cover-type changes.

To address the problem of mixed pixel within the remote sensing community, a superresolution technique long studied by the computer science community has been proposed. Until now, hundreds of superresolution approaches have been proposed, which can be grouped into three categories: interpolation-based [25], construction-based [26], and learning-based [27]. The sparse learning-based superresolution methods outperformed the others and were recognized as an outstanding representative of the learning-based approach. The original one was developed by Yang et al. [28], in which a pair of dictionaries was first learned from prior data, and then applied for downscaling the CR data. Huang et al. [29] extended this approach for spatial and temporal data fusion, and experimental results also show that it outperforms other spatial and temporal data-fusion approaches when compared with actual observations with respect to reflectance fidelity. However, its suitability for actual LCC detection has not been tested. To make use of multisource data fusion for change detection, recent works investigated the use of prior land cover products for generating better change detection results [30,31]. Finally, it is worth mentioning the work in [32],

in which a learning-based approach was investigated to allow the achievement of high sub-pixel forest mapping accuracy.

In this study, a novel learning-based approach will be presented to detect LCCs at finer spatial resolution using multitemporal CR data. The proposed approach has two advantages. First, it is well designed to learn the LCC dynamics from previous multisource multitemporal satellite data directly, which indicates that the trained detector has a high capability in detecting high-quality LCC using similar but CR satellite data. Second, the proposed approach makes use of the finer land cover product to provide rich spatial details within a coarse pixel.

The remainder of this paper is organized as follows. In Section 2, the theoretical background and the proposed approach are fully introduced. In Section 3, fused results are validated and applied for LCC detection with actual images in the Guangzhou study area, China. The discussion and conclusions are given in Sections 4 and 5, respectively.

## 2. Materials and Methods

The proposed approach includes two main steps. First, the CR satellite data at the predicted time ( $t_1$ ) coupled with pairs of coarse and finer resolution satellite data at previous times (e.g.,  $t_0$ ) were used to produce a finer resolution synthetic data at the predicted time ( $t_1$ ). Second, the LCC was detected at finer spatial resolution using the obtained finer resolution synthetic data and previous land cover maps.

### 2.1. Learning-Based Approach for Generating Finer Resolution Synthetic Satellite Data

It is an extremely ill-posed problem to infer the HR data directly from CR data. In this study, we will solve the problem from the perspective of LCC recovery. The recovered changed data were added with the high-resolution satellite data at a previous time to obtain the final downsampled image at the predicted time. Under a mild condition, it can be assumed that actual LCC from bitemporal satellite images can be sparsely represented as a linear combination of different LCC bases. As the following shows, a high-resolution LCC patch can be represented as a linear combination of LCC patterns with respect to a dictionary.

$$\Delta X \approx D_h \alpha \text{ where } : |\alpha|_0 \leq K \quad (1)$$

where  $\Delta X$  is an LCC patch with HR,  $D_h$  is a high-resolution dictionary,  $\alpha$  is the sparse representation coefficient, and  $K$  is the number of bases for the dictionary  $D_h$ .

It is further assumed that a high-resolution LCC patch can be degraded into a CR LCC patch with respect to a projection matrix. Then, the degraded CR LCC patch can also be inferred to have sparse representations with respect to a low-resolution dictionary, as the following formula shows:

$$\Delta Y \approx A \Delta X = A D_h \alpha = D_l \alpha \text{ where } : |\alpha|_0 \leq K \quad (2)$$

where  $\Delta Y$  and  $\Delta X$  are CR and HR LCC patches, respectively,  $A$  represents the projection matrix from  $\Delta X$  to  $\Delta Y$ ,  $D_h$  and  $D_l$  are a pair of dictionaries, and  $\alpha$  is the estimated coefficient.

Both HR and CR patches have the same sparse representations, and their co-occurrence can be captured by using a pair of coupled dictionaries. Thus, the downscaling issue for estimating HR data from CR data can be transformed into another issue, where both sparse coefficient and coupled dictionaries need to be estimated. There are two main steps to achieving this target: (1) dictionary learning—a pair of dictionaries was learned by using sample patch data from a pair of CR and HR data, of which each column (base) represents a specific LCC pattern; (2) sparse representation—sparse coefficients are estimated to reconstruct the HR LCC from CR LCC patches. The details are given below.

### 2.1.1. Dictionary Learning

Because the individual sparse coding problem of LCC in the high-resolution and low-resolution patches can be represented by the sparse linear combinations with respect to  $D_h$  or  $D_l$ , these two targets (see Formulas (1) and (2)) can be combined to form a unique target as shown below:

$$\min_{\{D_h, D_l, \alpha\}} \|\Delta X - D_h \alpha\|_2^2 + \|\Delta Y - D_l \alpha\|_2^2 + \gamma \|\alpha\|_1 \quad \text{where : } |\alpha|_0 \leq K \quad (3)$$

where  $\Delta X$  is a change patch for HR data,  $D_h$  is the HR dictionary,  $\alpha$  is the sparse representation coefficient, and  $K$  is the number of bases for dictionary  $D_h$ .

With the same learning strategy as Yang et al. [28], sampled training image patch pairs are first sampled from previously acquired low- and high-resolution data before a pair of dictionaries is jointly trained with these sampled patches using the  $k$ -singular value decomposition algorithm [33].

### 2.1.2. Sparse Representation

Sparse representation was then used to estimate the sparse coefficient and finally recover the HR LCC data. Based on the sparse representation of HR image patch shown in Formula (1), the solution of the sparse coefficients of a specific HR patch ( $\Delta X_s$ ) can be obtained via the following optimization function:

$$\Delta X_s = D_h \times \alpha^* \quad (4)$$

$$\text{where } \alpha^* : \min \|\alpha\|_1 \text{ s.t. } \begin{cases} \|D_l \alpha - \Delta Y_s\|_2^2 \leq \varepsilon_1 \\ \|D_h \alpha - W\|_2^2 \leq \varepsilon_2 \end{cases} \quad (5)$$

where  $\Delta X_s$  is a change patch for HR data at location  $s$ ,  $D_h$  and  $D_l$  are trained dictionaries for both HR and CR LCC,  $\alpha$  is the sparse coefficient that needs to be estimated, and  $W$  is the overlap between the current target patch and the previously reconstructed high-resolution patch. As recommended in [27,29], the dictionary size used in this study was set to 256, and the patch size was set to  $8 \times 8$ .

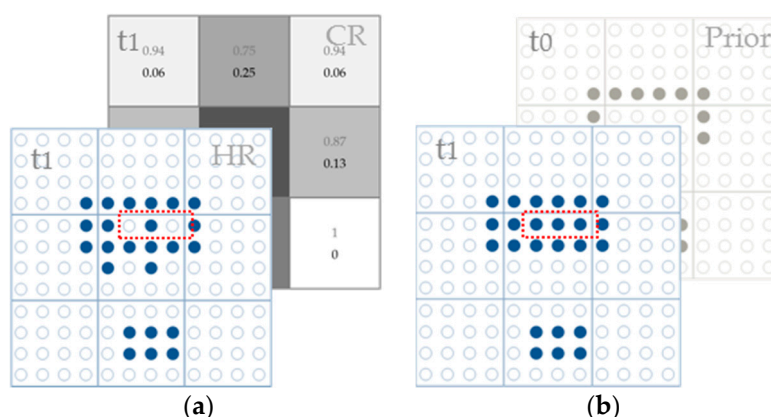
The process is operated patch by patch. If the sparse coefficient for each patch is sufficiently sparse, HR LCCs should then be recovered from the patches of CR LCCs with respect to the trained dictionaries. To agree with the previously computed adjacent high-resolution patches, a balance term (seen in the second term of Formula (3)) was used to preserve the fidelity of previous recovered LCC patches. Once the sparse coefficient is estimated, then the finer LCC patch can be recovered with Formula (4). Herein, the orthogonal matching pursuit algorithm was used to estimate the sparse coefficient [34].

The above procedure can be used to estimate the HR LCCs. Finally, the recovered LCCs are added with the HR image at the previous time to obtain the final downscaled image at the predicted time.

## 2.2. Sub-Pixel Change Detection with Synthetic Satellite Data

In the following, the obtained synthetic satellite data with a finer spatial resolution coupled with the land cover product at a previous time were used to detect LCCs at a finer spatial resolution. Given that the synthetic satellite data are not the real satellite data at the predicted time, the land cover map from the synthetic Landsat data appears to be different from actual land cover patterns. Figure 1a shows the initial land cover map obtained from synthetic data, and it appears to have some incorrect classification results at a sub-pixel level (highlighted with red). Thus, in this step, the obtained land cover map needs to be adjusted to ensure that it is consistent with prior land cover patterns as well as finer land cover products at previous times. The change detection procedure involves the following three steps. For more details, refer to [30].

First, a land cover map at the predicted time was produced from the obtained synthetic data using a supervised classification method, then land cover proportions at a CR were estimated from the obtained finer resolution land cover map.



**Figure 1.** Illustration of precise land cover mapping using synthetic Landsat data. (a) Synthetic Landsat data and its land cover proportions; (b) Land cover map at the predicted time (t1) using the proposed approach.

Second, based on the obtained land cover proportions, sub-pixel labels are initially randomly allocated maintaining the proportions. After random initialization, the labels of sub-pixels are iteratively swapped by counting their spatial correlations with surrounding pixels, and finally, the labels of sub-pixels are consistent with their neighborhood. The surrounding pixels include the nearby pixels at the current predicted time as well as neighboring pixels from land cover maps at previous times. For example, Figure 1a shows the initial land cover map obtained from synthetic data and its land cover proportions, while Figure 1b shows the final obtained land cover map at the predicted time (t1) using both the land cover proportions and a finer resolution land cover product at a previous time (t0).

Third, a refined land cover map at the predicted time was achieved via the above two steps. A change detection result (t1–t0) can be made by comparing the land cover map at the predicted time (t1) with the map from a previous time (t0).

### 3. Experiments and Result Analysis

The proposed approach was tested using actual data in the study area of Guangzhou, China (23°N, 113°E). This area has experienced a high percentage of land use/LCC during the past several decades, where most of the farmlands and forestlands have been changed into built-up areas due to rapid urbanization. The accurate monitoring of its rapid LCC is beneficial for the scientific management and sustainable development of this area.

Three pairs of medium-resolution Landsat and CR MODIS data for 31 October 2000, 7 November 2002, and 3 October 2004 were acquired for this study area. In this study, the MODIS reflectance products (MOD09GA) provided by NASA were adopted, and these products have been atmospherically corrected to land surface reflectance. For the original Landsat-5 data, they were atmospherically corrected into land surface reflectance using the atmospheric correction tool FLAASH [20]. Moreover, the downloaded MODIS data products were geometrically corrected to the same geographical area as the Landsat data, so both the MODIS and Landsat data cover the same extent. Based on the acquired satellite data, the preprocessed pairs of Landsat and MODIS data for the years 2000 and 2002 were used as training data, while the actual Landsat data for 2004 were used as validation data.

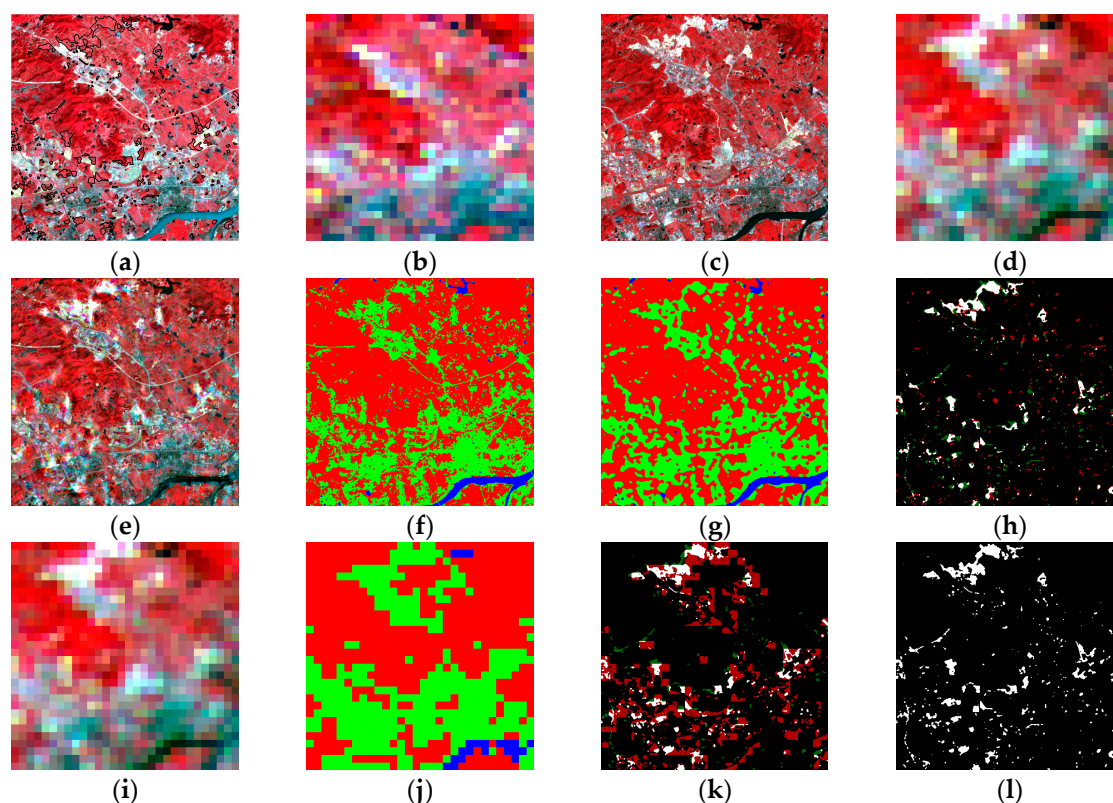
#### 3.1. Synthetic Data Generation and Sub-Pixel Change Detection

Synthetic Landsat-like data for 2004 were predicted via the following main steps. First, some low- and high-resolution LCC patches were randomly sampled from the achieved difference image to



train the dictionary, while the difference image reflects LCC from year 2000 to 2002 with the acquired satellite data at years 2000 and 2002. Next, the sparse learning approach introduced in the above section was used to recover HR LCCs from year 2002 to 2004 with respect to the coarse difference image and a pair of dictionaries. Finally, high-quality synthetic Landsat data at the predicted time were recovered by adding the predicted high-resolution difference data to previous Landsat data.

Based on a pair of Landsat and MODIS satellite data for the year 2002 (shown in Figure 2a,b) and MODIS data for the year 2004 (shown in Figure 2d), the finer resolution Landsat-like data for 2004 using the proposed learning-based approach are given in Figure 2e. Using the synthetic Landsat data for 2004, two sets of land cover maps (including the initial and final ones) were generated and are shown in Figure 2f,g. To validate its performance in detecting LCC from year 2002 to 2004, the synthetic Landsat data at year 2004 coupled with the prior land cover product from 2002 were used to generate an LCC map from 2002 to 2004 (Figure 2h). It shows the change detection result using the synthetic satellite data, in which white was used to reflect the correctly predicted LCC classes. In comparison, the MODIS data at year 2004 were also used to generate a change detection result (shown in Figure 2k based on the land cover map at CR (shown in Figure 2j)). The actual LCC map from year 2002 to 2004 provided in Figure 2l was used for validation.



**Figure 2.** Test with the actual MODIS data for sub-pixel change detection using different downscaling methods: (a) Landsat data for the year 2002 and the actual LCC from 2002 to 2004 (highlighted with black); (b) MODIS data for 2002; (c) Landsat for 2004 as a reference; (d) MODIS data for 2004; (e) Fused result for 2004 with the proposed approach; (f) Initial land cover map for 2004 with the fused result shown in Figure 1e; (g) Final land cover map for 2004 with the initial land cover map using the proposed approach; (h) Change detection result using the proposed approach; (i) MODIS data for 2004; (j) Land cover map for 2004 with MODIS data; (k) Change detection result with MODIS data from 2002 to 2004; (l) Actual LCC from 2002 to 2004 for validation.

### 3.2. Accuracy Assessment

In the following, the accuracy of different change detection results using different approaches was assessed. Five different scalars—namely, the Kappa statistic, the overall accuracy (OA), the commission error (CE), the omission error (OE), and the correlation coefficient (CC)—were used to assess the change detection accuracy. Other than the omission and commission errors, a higher value of each index reflects a higher change detection accuracy.

Change detection accuracy statistics of the fused results using different approaches are given in Table 1. Results using the simulated MODIS data are also provided for comparison, as shown on the right side of Table 1. It is found that the change detection accuracy with the fused result is much better than using the original MODIS data. Moreover, the proposed approach gives slightly better results than the STARFM method for all scale factors used.

**Table 1.** Change detection accuracy for the fused result with different methods. STARFM: spatial and temporal adaptive reflectance data-fusion model.

|       | Actual Data |        |          | Simulated Data (S = 16) |        |          |
|-------|-------------|--------|----------|-------------------------|--------|----------|
|       | Soft        | STARFM | Proposed | Soft                    | STARFM | Proposed |
| Kappa | 0.45        | 0.46   | 0.47     | 0.46                    | 0.49   | 0.50     |
| OA    | 83%         | 84%    | 85%      | 83%                     | 85%    | 86%      |
| CE    | 19%         | 18%    | 17%      | 18%                     | 17%    | 17%      |
| OE    | 32%         | 38%    | 37%      | 31%                     | 36%    | 30%      |
| CC    | 0.68        | 0.69   | 0.78     | 0.77                    | 0.86   | 0.89     |

## 4. Discussion

### 4.1. Strengths

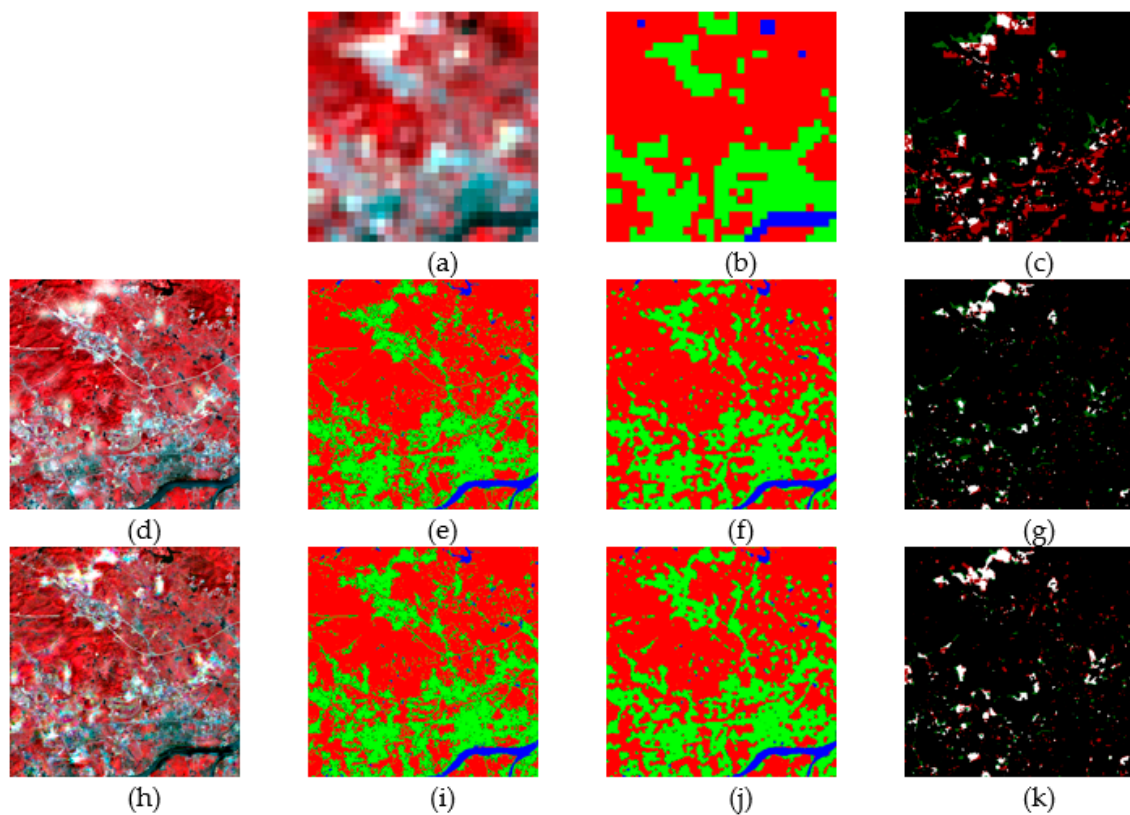
It is apparent that the fusion-based approaches—including STARFM and the proposed one—perform better than the soft classification method when CR data are directly used based on the accuracy statistics provided in Table 1. Let's take the simulated satellite data as an example. Overall accuracies for the results with the proposed and STARFM methods are 86% and 85%, respectively. Varying 83% for the soft classification method is also obtained. Especially, the soft classification approach tends to overestimate the actual LCC, while the fusion-based approach can improve it. When the two downscaling approaches are compared with each other, it is found that the learning-based approach performs slightly better than the conventional STARFM method in terms of all tested indices. Results with the proposed approach, moreover, tend to have better CCs than STARFM. The advantage of the proposed approach is that it can learn the change pattern or spatial texture information from previous satellite data, which is better than the STARFM.

Compared with the conventional unmixing-based fusion approach, an optimized neighborhood size is not required for the proposed approach. Because it is possible to achieve the desired result with a patch covering the whole area of a coarse pixel, a patch size of  $8 \times 8$  Landsat pixels was used in this study. In addition, the number of land cover types is not required, as a large number of bases (chosen as 256 in this study) is enough to reflect the whole LCC patterns in this study.

### 4.2. Scale Effect

To assess the impact of spatial scale of the proposed approach in monitoring LCC information, a series of simulated data degraded from the actual Landsat data were used in this study. In our experiment, the scale factors of 4, 8, and 16 were tested. Based on the simulated MODIS and Landsat data, finer resolution fused satellite data and land cover/change maps at different scales were generated. Figure 3h–k shows the results using the proposed approach, while Figure 3a–g show the results using the original MODIS data and the conventional STARFM method for comparison,

respectively. Accuracy statistics for all mapping results with different approaches are provided in Table 2.



**Figure 3.** Sub-pixel change detection results with the simulated MODIS data using different methods at a scaling factor of 16. The upper row shows the results using simulated MODIS data ( $s = 16$ ): (a) Simulated MODIS data for the year 2004; (b) Land cover map using simulated MODIS data; and (c) Change detection result using simulated MODIS data. The middle row shows the results using the conventional fusion-based method. (d) Synthetic Landsat data for 2004 using the STARFM method; (e) Initial land cover map from the result shown in (d); (f) Final land cover map from synthetic Landsat data using the STARFM method; and (g) Change detection result from 2002 to 2004 using the STARFM method. The lower row shows the results using the proposed approach. (h) Synthetic Landsat data for 2004 using the proposed approach; (i) Initial land cover map from the result shown in (h); (j) Final land cover map using the proposed approach; (k) Change detection result from 2002 to 2004 using the proposed approach.

**Table 2.** Change detection accuracy for the fused result with different methods under different scale factors. OA: overall accuracy; CE: commission error; OE: omission error; CC: correlation coefficient.

|       | Scale Factor = 4 |        |          | Scale Factor = 8 |        |          | Scale Factor = 16 |        |          |
|-------|------------------|--------|----------|------------------|--------|----------|-------------------|--------|----------|
|       | Soft             | STARFM | Proposed | Soft             | STARFM | Proposed | Soft              | STARFM | Proposed |
| Kappa | 0.53             | 0.60   | 0.61     | 0.52             | 0.53   | 0.55     | 0.46              | 0.49   | 0.50     |
| OA    | 86%              | 89%    | 90%      | 85%              | 87%    | 88%      | 83%               | 85%    | 86%      |
| CE    | 15%              | 13%    | 15%      | 16%              | 15%    | 15%      | 18%               | 17%    | 17%      |
| OE    | 23%              | 27%    | 16%      | 23%              | 30%    | 26%      | 31%               | 36%    | 30%      |
| CC    | 0.89             | 0.92   | 0.93     | 0.88             | 0.89   | 0.92     | 0.77              | 0.86   | 0.89     |

According to the accuracy statistics provided in Table 2, three observations can be summarized as below. First, the change detection accuracy decreases significantly as the scale factor increases. Taking the results using the proposed approach as an example, the Kappa index decreased from 0.61



to 0.50 when the scale factor increased from 4 to 16. Second, when the performances of different approaches were compared, it was found that the fused-based approaches achieved better change detection accuracy than results using the original MODIS data. In particular, the use of original MODIS data tended to overpredict the actual LCC, while the fused-based approaches improved it. Third, comparing the performance of STARFM and the proposed method, the proposed approach performed better than STARFM, regardless of which scale factor was used. In particular, a much better CC was achieved by the proposed learning-based approach compared with STARFM, indicating that the learning-based approach is suitable for the downscaling of CR data.

#### 4.3. Limitations

Although the proposed approach has been validated and proven suitable for sub-pixel LCC detection using CR satellite data, there are still some limitations. First, misregistration errors between multisource satellite data of the proposed approach may affect the final change detection results, and thus using the simulated data can achieve better detection accuracy than using actual satellite data. Second, the advantage of the proposed approach is obvious when the predicted LCC percentages are compared with others by referring to the CC index. Nevertheless, the predicted LCCs still have positional errors within a coarse pixel, which may offset its advantage for sub-pixel LCC detection using the actual multisource satellite data. Lastly, it is a computationally expensive approach. Both dictionary training and sparse coefficient estimation processes are computationally expensive.

### 5. Conclusions

In this paper, a learning-based downscaling method is presented to generate finer resolution LCC results using prior LCC information and one CR data at the predicted time, in which prior LCC patterns are learned and modeled using the popular sparse learning approach. Further experiments demonstrate that it is better than the conventional downscaling approach STARFM when both predicted synthetic data are applied for LCC detection. Experiments conducted at Guangzhou show that the proposed learning-based approach outperforms both the conventional change detection method and the fusion-based change detection method. According to the results with the proposed approach using actual MODIS data, the overall LCC detection accuracy is 85%, which is better than the results using a conventional soft classification method and fusion-based method (83% and 84%, respectively). More importantly, it is found that high-quality LCC percentages—as indicated by the CC index—can be achieved by the proposed approach, as the CC index for the proposed approach is 0.78, which is much better than the results using soft classification and fusion-based methods (0.68 and 0.69, respectively). This finding is meaningful for high-quality LCC detection at a sub-pixel level.

This study also investigated the effect of scale factor on sub-pixel change detection. In particular, results from fusion-based approaches perform much better than when the original coarse satellite data are used directly, regardless of the scale factor used. The soft classification method tends to overestimate the actual LCC area, while the proposed approach tends to miss some actual LCC area. When a large scale factor is adopted, the proposed approach performs slightly better than the STARFM model. When compared with the use of the simulated MODIS data, the use of actual MODIS data achieves a slightly lower LCC detection accuracy. The main reason may be due to the positioning error between MODIS and Landsat data, which needs to be further investigated.

**Acknowledgments:** This research was supported by the China NSFC project under contract 61373114, 61661166011, and 11690011, the VC's discretionary fund of The Chinese University of Hong Kong, Macau Science and Technology Development Funds with No. 003/2016/AFJ, and 973 Program of China with No. 2013CB329404.

**Author Contributions:** Y.X. and D.M. conceived the experiments and interpreted the result. Y.X. and L.L. performed the experiments. Y.X. wrote the paper.

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

## References

1. Bruzzone, L.; Prieto, D.F. An adaptive semi-parametric and context-based approach to unsupervised change detection in multitemporal remote-sensing images. *IEEE Trans. Image Process.* **2002**, *11*, 452–466.
2. Xu, Y.; Huang, B. Spatial and temporal classification of synthetic satellite imagery: Land cover mapping and accuracy validation. *Geo-Spat. Inf. Sci.* **2014**, *17*, 1–7. [[CrossRef](#)]
3. Singh, A. Digital change detection techniques using remotely sensed data. *Int. J. Remote Sens.* **1989**, *10*, 989–1003. [[CrossRef](#)]
4. Wang, Q.; Lin, J.; Yuan, Y. Salient band selection for hyperspectral image classification via manifold ranking. *IEEE Trans. Neural Netw. Learn. Syst.* **2016**, *27*, 1279–1289. [[CrossRef](#)] [[PubMed](#)]
5. Fung, T.; LeDrew, E. Application of principal components analysis to change detection. *Photogramm. Eng. Remote Sens.* **1988**, *53*, 1649–1658.
6. Chen, J.; Gong, P.; He, C.; Pu, R.; Shi, P. Land-use/land-cover change detection using improved change-vector analysis. *Photogramm. Eng. Remote Sens.* **2003**, *69*, 369–379. [[CrossRef](#)]
7. Lu, D.; Mausel, P.; Brondizio, E.; Moran, E. Change detection techniques. *Int. J. Remote Sens.* **2004**, *25*, 2365–2401. [[CrossRef](#)]
8. Le Hégat-Masclé, S.; Ottlé, C.; Guérin, C. Land cover change detection at coarse spatial scales based on iterative estimation and previous state information. *Remote Sens. Environ.* **2005**, *95*, 464–479. [[CrossRef](#)]
9. Strugnell, N.C.; Lucht, W.; Schaaf, C. A global albedo data set derived from AVHRR data for use in climate simulations. *Geophys. Res. Lett.* **2001**, *28*, 191–194. [[CrossRef](#)]
10. Friedl, M.A.; McIver, D.K.; Hodges, J.C.; Zhang, X.Y.; Muchoney, D.; Strahler, A.H.; Woodcock, C.E.; Gopal, S.; Schneider, A.; Cooper, A.; et al. Global land cover mapping from MODIS: Algorithms and early results. *Remote Sens. Environ.* **2002**, *83*, 287–302. [[CrossRef](#)]
11. Walker, J.J.; De Beurs, K.M.; Wynne, R.H.; Gao, F. Evaluation of Landsat and MODIS data fusion products for analysis of dryland forest phenology. *Remote Sens. Environ.* **2012**, *117*, 381–393. [[CrossRef](#)]
12. Verbesselt, J.; Hyndman, R.; Zeileis, A.; Culvenor, D. Phenological change detection while accounting for abrupt and gradual trends in satellite image time series. *Remote Sens. Environ.* **2010**, *114*, 2970–2980. [[CrossRef](#)]
13. Wu, K.; Du, Q.; Wang, Y.; Yang, Y. Supervised Sub-Pixel Mapping for Change Detection from Remotely Sensed Images with Different Resolutions. *Remote Sens.* **2017**, *9*, 284. [[CrossRef](#)]
14. He, D.; Zhong, Y.; Feng, R.; Zhang, L. Spatial-Temporal Sub-Pixel Mapping Based on Swarm Intelligence Theory. *Remote Sens.* **2016**, *8*, 894. [[CrossRef](#)]
15. Ling, F.; Li, W.; Du, Y.; Li, X. Land cover change mapping at the subpixel scale with different spatial-resolution remotely sensed imagery. *IEEE Geosci. Remote Sens. Lett.* **2011**, *8*, 182–186. [[CrossRef](#)]
16. Ling, F.; Foody, G.M.; Li, X.; Zhang, Y.; Du, Y. Assessing a temporal change strategy for sub-pixel land cover change mapping from multi-scale remote sensing imagery. *Remote Sens.* **2016**, *8*, 642. [[CrossRef](#)]
17. Zurita-Milla, R.; Kaiser, G.; Clevers, J.G.P.W.; Schneider, W.; Schaepman, M.E. Downscaling time series of MERIS full resolution data to monitor vegetation seasonal dynamics. *Remote Sens. Environ.* **2009**, *113*, 1874–1885. [[CrossRef](#)]
18. Foody, G.M. Approaches for the production and evaluation of fuzzy land cover classifications from remotely-sensed data. *Int. J. Remote Sens.* **1996**, *17*, 1317–1340. [[CrossRef](#)]
19. Atkinson, P.M. Sub-pixel target mapping from soft-classified, remotely sensed imagery. *Photogramm. Eng. Remote Sens.* **2005**, *71*, 839–846. [[CrossRef](#)]
20. Gao, F.; Masek, J.; Schwaller, M.; Hall, F. On the blending of the Landsat and MODIS surface reflectance: Predicting daily Landsat surface reflectance. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 2207–2218.
21. Hilker, T.; Wulder, M.A.; Coops, N.C.; Linke, J.; McDermid, G.; Masek, J.G.; Gao, F.; White, J.C. A new data fusion model for high spatial-and temporal-resolution mapping of forest disturbance based on Landsat and MODIS. *Remote Sens. Environ.* **2009**, *113*, 1613–1627. [[CrossRef](#)]
22. Hilker, T.; Wulder, M.A.; Coops, N.C.; Seitz, N.; White, J.C.; Gao, F.; Masek, J.G.; Stenhouse, G. Generation of dense time series synthetic Landsat data through data blending with MODIS using a spatial and temporal adaptive reflectance fusion model. *Remote Sens. Environ.* **2009**, *113*, 1988–1999. [[CrossRef](#)]
23. Zhu, X.; Chen, J.; Gao, F.; Chen, X.; Masek, J.G. An enhanced spatial and temporal adaptive reflectance fusion model for complex heterogeneous regions. *Remote Sens. Environ.* **2010**, *114*, 2610–2623. [[CrossRef](#)]

24. Roy, D.P.; Ju, J.; Lewis, P.; Schaaf, C.; Gao, F.; Hansen, M.; Lindquist, E. Multi-temporal MODIS–Landsat data fusion for relative radiometric normalization, gap filling, and prediction of Landsat data. *Remote Sens. Environ.* **2008**, *112*, 3112–3130. [[CrossRef](#)]
25. Sun, J.; Xu, Z.; Shum, H.Y. Image super-resolution using gradient profile prior. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Anchorage, AK, USA, 23–28 June 2008; IEEE: Hoboken, NJ, USA, 2008.
26. Baker, S.; Kanade, T. Limits on super-resolution and how to break them. *IEEE Trans. Pattern Anal. Mach. Intell.* **2002**, *24*, 1167–1183. [[CrossRef](#)]
27. Gu, S.; Zuo, W.; Xie, Q.; Meng, D.; Feng, X.; Zhang, L. Convolutional sparse coding for image super-resolution. In Proceedings of the IEEE International Conference on Computer Vision, Santiago, Chile, 13–16 December 2015; pp. 1823–1831.
28. Yang, J.; Wright, J.; Huang, T.S.; Ma, Y. Image super-resolution via sparse representation. *IEEE Trans. Image Process.* **2010**, *19*, 2861–2873. [[CrossRef](#)] [[PubMed](#)]
29. Huang, B.; Song, H. Spatiotemporal reflectance fusion via sparse representation. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 3707–3716. [[CrossRef](#)]
30. Xu, Y.; Huang, B. A spatio-temporal pixel-swapping algorithm for subpixel land cover mapping. *IEEE Geosci. Remote Sens. Lett.* **2014**, *11*, 474–478. [[CrossRef](#)]
31. Wang, Q.; Shi, W.; Atkinson, P.M.; Li, Z. Land cover change detection at subpixel resolution with a Hopfield neural network. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, *8*, 1339–1352. [[CrossRef](#)]
32. Zhang, Y.; Atkinson, P.M.; Li, X.; Ling, F.; Wang, Q.; Du, Y. Learning-Based Spatial–Temporal Superresolution Mapping of Forest Cover with MODIS Images. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 600–614. [[CrossRef](#)]
33. Aharon, M.; Elad, M.; Bruckstein, A. K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Trans. Signal Process.* **2006**, *54*, 4311–4322. [[CrossRef](#)]
34. Davis, G.; Mallat, S.; Avellaneda, M. Adaptive greedy approximations. *Construc. Approx.* **1997**, *13*, 57–98. [[CrossRef](#)]



© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).