Maximizing Urban Features Extraction from Multi-sensor Data with
Dempster-Shafer Theory and HSI Data Fusion Techniques

Mohammed Oludare Idrees1*, Vahideh Saeidi1, Biswajeet Pradhan2, Helmi Zulhaidi Mohd Shafri2

1Department of Civil Engineering, Faculty of Engineering, Universiti Putra Malaysia, 43400 UPM, Serdang, Selangor Darul-Ehsan, Malaysia
2 Geospatial information Science Research Center (GISRC), Faculty of Engineering Universiti Putra Malaysia. 43400 UPM, Serdang, Selangor Darul-Ehsan, Malaysia

*Corresponding author’s email: dare.idrees[AT]gmail.com

ABSTRACT—This paper compares two multi-sensor data fusion techniques – Dempster-Sharfer Theory (DST) and Hue Saturation Intensity (HSI). The objective is to evaluate the effectiveness of the methods in term of space and time and quality of information extraction. LiDAR and hyperspectral data were fused using the two methods to extract urban landscape features. First, digital surface model (DSM), LiDAR intensity and hyperspectral image were fused with HSI. Then the result was classified into five classes (metal roof building, non-metal roof building, tree, grass and road) using supervised classification (minimum distance) and the classification accuracy assessment was done. Second, Dempster Shafer Theory (DST) utilized the evidences available to fuse normalized DSM, LiDAR intensity and hyperspectral derivatives to classify the surface materials into five classes as before. It was found out that DST perform well in the ability to discriminate different classes without expert information from the scene. Overall accuracy of 87% achieved using DST. While in HSI technique, the overall accuracy obtained was 74.3%. Also, metal and non-metal roof types were clearly classified with DST which, does not have a good result with HSI. A fundamental setback of HSI is its limitation to fusion of only two sensor data at a time whereas we could integrate different sensor data with DST. Besides, the time required to select training site for supervised classification, the accuracy of feature classification with HSI fused data is dependent on the knowledge of the analyst about the scene with the other one. This study shows DST to be an accurate and fast method to extract urban features and roof types. It is hoped that the increasing number of remote sensing technology transforming to era of redundant data will make DST a desired technique available in most commercial image processing software packages.

Keywords—Data fusion, Feature extraction, urban mapping, Hyperspectral, LiDAR

1. INTRODUCTION

The global surge in rural-urban migration in search of better live, employment opportunities, and education exerts pressure on the population of urban environment leading to saturation of small spaces with various artificial surface materials. According to United Nations Economics and Social Affairs report, as at 2010, Malaysia has a total population of 28,250,458 [1], out of which it is estimated that 73% will be living in urban areas by 2013 [2]. This factor has brought a lot of burden and congestion on the urban environment that necessitates effective planning to cope with the resultant impact on the socioeconomic and environmental challenges.

Multispectral images such as IKONOS, QuickBird, AVHRR, SPOT, etc. have been widely used for land surface cover inventory of urban environment [3], [4]. However, limitation of spectral resolution render them inadequate to obtain detailed analysis of surface material within urban area [3], [5], [6]. In contrast to the multispectral sensors that collect reflected energy from surface material in discrete wavelength of usually less than 10 bands, hyperspectral sensors record spectral signature of surface material in dozens or hundreds as a series of narrow and continuous wavelength bands [7]. Advances in hyperspectral remote sensors give rise to new prospect for detailed inventory of surface material, particularly for geological, forestry, agricultural, and urban applications [8], [9]. Although hyperspectral is highly useful for diagnostic mapping, particularly for identifying the internal structure and/or chemical properties of material, cost of data collection and limited spatial coverage are a major drawback.

There are several studies on hyperspectral for urban mapping [3], [4], [7], [10], [11] but the absence of height information create a void in some applications such as roof structure and material. Essentially, roof had been, from time immemorial, a shield that protects people and their properties against weather and other external threat [5]. Nonetheless, increasing concern for health and environmental sustainability have brings the use of roof covers made of environmentally friendly material into focus. This concerns range from assessment of pollutants on urban surfaces, rainwater, runoff, and underground water contamination which are dependent on slope and exposition of roof material.
urban heat island effects and thermal performance of building [13], heat conservation and energy loss [6],
development of solar generating power and building exterior planning [14]. The urban environment is comprised of
complex geometric structure and heterogeneous surface materials. So, to optimally extract complete information,
combination of high resolution data with respect to spatial geometry and spectral discrimination is important. Therefore,
the fusion of hyperspectral with 3D data such as SAR [15] and LiDAR [16] data for urban information extraction has
gain preference. [17], [9], [12], and [18] are some of the successful works done to map urban features by combining
hyperspectral and LiDAR data.

Figure 1: Metal roofing material with solar panel installed (Source: Englert, 2013)

Today, the market is flooded with varieties of roofing materials ranging from tiles, bricks, fibred-cement sheet,
corrugated metal roofing sheet, polycarbonate, textile-fibre, asbestos, clay roofing materials, etc. Figure 1 shows a metal
roofing material with solar panel installed on it. With the concept of sustainability in focus, metal roof is receiving new
attention because of its advantages that include wind and fire resistance, saving energy inside the homes and low weight
and long lasting to compare with other materials. The list of benefits offered by metal roof can be found in [19] and [14].
The objective of this study is to extract urban features from fusion of hyperspectral and LiDAR data, classify building
roof cover into metal and non-metal, and compare the performance of DST and HSI fusion techniques.

2. METHOD AND RESULTS

2.1 Study area and data description

The Advanced Imaging Spectrometer for Applications (AISA) and TerraScan LiDAR data collected over Universiti
Putra Malaysia (UPM) campus were used for this study. Figure 2 shows the study area within the campus extracted from
Google Earth. The hyperspectral image consists of 20 spectral bands with a spectral range of 400-970nm in the visible
and near-infrared spectrum of the electromagnetic energy with a spatial resolution of 1m and spectral resolution of
2.9nm. The image was acquired in 2004 by Aeroscan Precision (M) Sdn Bhd. Also, the LiDAR data was acquired in
2009 using Riegl LiDAR sensor and cannon III camera on-board aircraft flying at an altitude of approximately 5000m
above MSL with average point density of 2.3 points/m².
Figure 2: Study area with image extracted from Google Earth

Figure 3 presents the methodological work flow of the entire procedure. Two fusion methods were implemented for urban feature classification. First, hyperspectral and DSM were fused using HSI technique and the result classified into five land cover classes. Second, hyperspectral image and nDSM generated from LiDAR data were fused and classified into five land cover types using DST. At the final stage, classification accuracy of both methods were assessed and compared.

Figure 3: Methodological work flow of the two fusion techniques
2.2 Data Preprocessing

Hyperspectral image processing was done in ENVI 4.8. Preprocessing commenced with atmospheric correction using the internal average relative reflectance (IARR). Thereafter, the atmospherically corrected hyperspectral image was registered to LiDAR DSM. With these done, the data was ready for further processes. For the LiDAR data, at first, DSM and DTM were generated from the point cloud using BCAL LiDAR (ENVI Toolkit) and resampled to 1m resolution. DSM includes not only the bare earth, but also the elevation of all other features above the earth’s surface, whereas DTM contains only the elevation of the terrain. The DSM was normalized by subtracting DTM from it to produce nDSM (called normalized DSM). According to [20] nDSM contains elevation of all objects above the ground (e.g. building and tree) and therefore a useful clue for identifying buildings and trees.

2.3 Data Fusion

2.3.1 Hue Saturation Intensity (HSI)

[21] Give details of the popular fusion algorithms, among which are the HSI used in this study. HSI algorithm transforms color images from the RGB (red-green-blue) into HSI color space [22], [23]. The intensity (I) band defines the total sensitivity (brightness) of the scene to all visible colors and ultraviolet light. The band is represented as a series of shades from white to black, thus, looks like a panchromatic image. The HSI fusion processes replace the intensity band with a high-resolution image (i.e. panchromatic image, DSM) and perform a reverse transformation of the new intensity band (DSM and LiDAR intensity in this case) with the hue and saturation bands producing the HSI fused image. Hue expresses the more important wavelength of the pixel and saturation measures the color concentration. The advantages of HSI is that it is easy to implement, and also, improves spatial resolution of the fused data [21], [22]. Meanwhile, spectral distortion which hampers the use of conventional classification methods and the restriction of the technique to only three or less images at a time are disadvantages [24].

Because of the above mentioned constraint for HSI technique, minimum noise fraction (MNF) transform and optimum index factor (OIF) were performed to reduce dimensionality of hyperspectral data and select three optimal bands. The three selected bands were fused with DSM as presented in figure 4.

![Figure 4: HSI input and output data: (a) Hyperspectral image (bands 19,11,2), ( b) DSM (c) Fused image](image-url)
2.3.2 Classification

HSI fused data was classified in five land cover types: metallic roof buildings, non-metal roof building, pavement, trees, and grass using supervised classification (minimum distance algorithm). Figure 5 shows the classified HSI fused data.

![Figure 5: Land cover classification map using HSI fusion](image)

Usually for supervised classification, prior knowledge of the study area (especially for metallic and non-metal roof types), adequate definition of land cover classes, and careful selection of training samples are required. However, selecting training site from HSI fused data for the purpose of classification suffers from spectral distortion which makes it difficult to discriminate some features like grass, bare earth and road as they exhibit similar properties such as appearance and height after fusion (figure 3). Moreover, classifying HSI fused data involves trial-and-error to identify appropriate method. That is to say, the classification accuracy is highly dependent method used. The classified image was quantitatively assessed and an overall accuracy of 87.5% and Kappa coefficient of 0.83 were obtained.

2.3.3 Dempster Shafer Theory (DST)

DST is a soft classifier [25] that mathematically combines different datasets and calculates associated evidential weight and reasoning to derive features through the measurement of support, uncertainty, plausibility and conflict [26], [27]. The uniqueness of DST is that, it can handle incomplete data coverage [28], and take into account, not only singleton classes (e.g. Building, Tree, Grass, and Road), but also union of the classes (e.g. BUT, BUTG and so on). Further details on DST can be found in [20], [29] and [30].

For five hypothetical land cover classes in , with a finite framework of exponent of 2 \(2^5\) consisting of singleton and union classes that can be generate from DST, only 12 classes provide valid evidences in nDSM, NDVI, band 8 and 14 used for this fusion. To do this, foremost, the mass function \(m(A)\) was defined using equation 1 to derive the probability in single class or union of classes.

\[
m(\emptyset) = \begin{cases} 0 & m(\emptyset) \leq 1 \\
\sum_{C \in \Theta} m(C) = 1 \\
m(\emptyset) = 0 \end{cases}
\]

where \(\emptyset\) interprets the empty class.
Then, for every class \( C \in 2^g \), support and plausibility were computed from mass functions [28]. At the final stage, rule of combination using equation 2 assigns multiple mass functions from individual contributor to a corresponding hypothetical class [25].

\[
\text{Combine}(C) = \left\{ \begin{array}{ll}
\frac{\sum_{C_k \cap C_j \neq \emptyset} \Pi_{j \in S} P_m(C_j)}{1-k}
\end{array} \right.
\]

where \( 1 - k \) is the measure of conflict in the cues

Finally, each pixel was classified into the greatest support class [31]. From the evidences in nDSM, it is possible to discriminate objects above the ground comprising of \( mR \cup nR \cup T \) with \( P_{\Delta H} \) (Table 1) mass function for the class, while \( G \cup R \) with probability value \( 1 - P_{\Delta H} \) complements the classes.

2.3.4 DST Reasoning

The probability mass functions modeling provides the apparent evidence of each class or union of the classes in each input dataset. There are some parameters and rules which should be defined prior to modeling the probability masses [32]. As an example, if within a dataset, it is possible to discriminate between \( C_1 \) and \( C_2 \) features where \( C_1 \cap C_2 = \emptyset \) and \( C_1 \cup C_2 = \emptyset \), then the probability mass can be defined as \( P_i(x) \) for each dataset \( i \) and each pixel \( x_j \) in the dataset using the parameters in equation 3.

\[
P_i(x) = \begin{cases} 
  P_1, & \forall x | x \leq x_1 \\
  P_2, & \forall x | x \geq x_2 
\end{cases}
\]

where \( 0 \leq P_1 \leq 1 \leq P_2 \) are probability masses to assign every pixel \( x_j \) using a predefined thresholds \( x_1 \) and \( x_2 \) while \( x_1 < x_2 \).

Table 1: Probability mass function for four datasets (nDSM, NDVI, band 8, and band 14)

<table>
<thead>
<tr>
<th>( C_j ) Class</th>
<th>nDSM ((\Delta H))</th>
<th>NDVI</th>
<th>Band8</th>
<th>Band14</th>
<th>Combined probability mass</th>
</tr>
</thead>
<tbody>
<tr>
<td>metallic Roof (mR)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>( P_{\Delta H} \times (1 - P_{NDVI}) \times (1 - P_{BG}) \times (1 - P_{B14}) )</td>
</tr>
<tr>
<td>non-metal Roof (nR)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>( P_{\Delta H} \times (1 - P_{NDVI}) \times (1 - P_{BG}) \times P_{B14} )</td>
</tr>
<tr>
<td>Tree (T)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>( P_{\Delta H} \times (1 - P_{NDVI}) \times (1 - P_{BG}) \times (1 - P_{B14}) )</td>
</tr>
<tr>
<td>Grass (G)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>( (1 - P_{\Delta H}) \times P_{NDVI} \times (1 - P_{BG}) \times (1 - P_{B14}) )</td>
</tr>
<tr>
<td>Road (R)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>( (1 - P_{\Delta H}) \times (1 - P_{NDVI}) \times (1 - P_{BG}) \times P_{B14} )</td>
</tr>
<tr>
<td>( G \cup T )</td>
<td>0</td>
<td>( P_{NDVI} )</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>( mR \cup nR \cup R )</td>
<td>0</td>
<td>1 - ( P_{NDVI} )</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>( mR \cup nR \cup T )</td>
<td>( P_{\Delta H} )</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>( G \cup R )</td>
<td>1 - ( P_{\Delta H} )</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>( nR \cup G \cup T \cup R )</td>
<td>0</td>
<td>0</td>
<td>1 - ( P_{BG} )</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
From Equation 2:

\[
1 - k = 1 - (1 - P_{n3H}) \times (1 - P_{NDVI}) \times P_{g9} \times P_{B14} + P_{n3H} \times (1 - P_{NDVI}) \times P_{g9} \times P_{B14} - P_{n3H} \times P_{NDVI} \times P_{g9} \times (1 - P_{B14}) - (1 - P_{n3H}) \times P_{NDVI} \times P_{g9} \times (1 - P_{B14})
\]

For instance, in NDVI, if the pixel value is greater than \(k\), the probability mass \(P_2\) (an index for support to be vegetation) will be assigned to the pixel. This shows the degree of membership of pixel in class \(C_4\). In the visible portion of electromagnetic spectrum, metallic roof (figure 6) has high reflectivity than other surface materials. So, identifying metal roof from one band in RGB with 671.875nm wavelengths can be a good clue to discriminate metal roof from other objects in the scene. Though, it is not assumed that all high reflective objects within that range are metallic properties. Hence, probability of 90% soft threshold was assigned. Likewise for band 14, more absorption of NIR occurs in Clay and tile material.

Figure 6: Hyperspectral signature (Source: http://auracle.ca/news/)

Therefore, hyperspectral band 14 with the wavelength of 740.760 nm was selected as a clue to discriminate nonmetal roof. Nevertheless, the existence of other objects such as road and pavement were taken into account (table 1). NDVI derived from hyperspectral data provides strong clue to the existence of vegetation (tree or grass) was therefore another dataset fed into the algorithm. The probability mass functions modeling and parameters for DST datasets are represented in table 2.

<table>
<thead>
<tr>
<th>X1</th>
<th>1.5m</th>
<th>50</th>
<th>150</th>
<th>50</th>
</tr>
</thead>
<tbody>
<tr>
<td>X2</td>
<td>3m</td>
<td>100</td>
<td>200</td>
<td>100</td>
</tr>
<tr>
<td>P1</td>
<td>0.05</td>
<td>0.08</td>
<td>0.10</td>
<td>0.90</td>
</tr>
<tr>
<td>P2</td>
<td>0.95</td>
<td>0.92</td>
<td>0.90</td>
<td>0.10</td>
</tr>
</tbody>
</table>

DST fuses nDSM, band 8, band 14, and NDVI to classify the five land cover types as presented in figure 7.
Figure 7: Landcover classification map using DST fusion

Similar to HSI method, the resulting classification accuracy assessment was done using ground truth and Google earth map with overall accuracy of 94.3% and Kappa coefficient of 0.93. Figure 8 shows the extracted metallic and nonmetallic roof types extracted from DST result overlaid on Google image of the study area.

Figure 8: Overlay of roof types on Google earth image of the study area
3. DISCUSSION

To compare the performance of the HSI and DST fusion techniques, the classification results were evaluated by calculating the matrix of confusion shown in Table 3.

Table 3: Classification accuracy assessment of HSI and DST

<table>
<thead>
<tr>
<th>Class</th>
<th>HSI Method</th>
<th>DST Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>non-metal Roof</td>
<td>89%</td>
<td>58%</td>
</tr>
<tr>
<td>Metallic Roof</td>
<td>77%</td>
<td>96%</td>
</tr>
<tr>
<td>Tree</td>
<td>64%</td>
<td>100%</td>
</tr>
<tr>
<td>Grass</td>
<td>99%</td>
<td>91%</td>
</tr>
<tr>
<td>Road</td>
<td>95%</td>
<td>96%</td>
</tr>
</tbody>
</table>

Overall Accuracy = 87.5%  Overall Accuracy = 94.3%
Kappa Coefficient = 0.83  Kappa Coefficient = 0.93

Quantitatively, both fusion methods have acceptable overall accuracy, however, DST shows better performance with overall accuracy 94.3% compared 87.5% for HSI and Kappa coefficient of 0.93 and 0.83 respectively. In addition, DST indicates wonderful discriminating ability for metallic material as presented in the table with producer accuracy of 98% against 77% for HSI, so also tree detection.

Similar to the quantitative analysis above, misclassification of metal to nonmetal roof is observed in the classified HSI fused data. For instance, the building to the right (upper corner) in figure 5 is a metal roof, as observed during the ground truth data collection (and confirmed in DST classified data), but classified as nonmetal in HSI. This may be due to distortion of spectral information, illumination or cloud shadow effect. Furthermore, it can be seen that all trees along paved surfaces are classified as building in both fusion techniques. The temporal variation between data acquisition date (2004 and 2009) contributes to this. During hyperspectral data collection, the tree crown and height were quite smaller to obstruct reflectance from pavement. However, in 2009 during LiDAR acquisition, the heights of the trees are much taller with wider canopy cover. This phenomenon is glaring in the fused image (figure 4c) where trees close to paved surfaces appear brighter as opposed to those with undergrowth grasses.

Table 4 presents the general comparison of the HSI and DST. DST implementation is not limited to a particular number of datasets or sensor types as long as evidence are available for spatial data modeling and reasoning [26]. This is not the case as with HSI which is limited to a finite maximum number of sensor types. Another distinguishing factor is the number of output that can be generated. DST algorithm can provide more information layers such as probability, plausibility, uncertainty, support, and conflict maps [30], whereas HSI can only generate a single data layer as output.

Table 4: Comparison between HSI and DST fusion techniques for land cover classification

<table>
<thead>
<tr>
<th>Fusion Method</th>
<th>HSI</th>
<th>DST</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of data input</td>
<td>Limited to 3</td>
<td>Not limited</td>
</tr>
<tr>
<td>Number of outcomes layer</td>
<td>1 layer</td>
<td>More than 1 layer</td>
</tr>
<tr>
<td>Expert knowledge from the scene</td>
<td>Needed</td>
<td>Not required</td>
</tr>
<tr>
<td>Band selection/reduction</td>
<td>Needed</td>
<td>Relies on evidence in individual</td>
</tr>
<tr>
<td>Classification algorithm</td>
<td>Trial and Error</td>
<td>Soft classification</td>
</tr>
<tr>
<td>Training Site</td>
<td>Needed for better Acc.</td>
<td>-</td>
</tr>
<tr>
<td>Spatial data modeling</td>
<td>-</td>
<td>Needed</td>
</tr>
<tr>
<td>Accuracy</td>
<td>-</td>
<td>Better</td>
</tr>
<tr>
<td>Availability in software</td>
<td>Common</td>
<td>Not common</td>
</tr>
</tbody>
</table>

Furthermore with DST, expert knowledge about the scene is not required but for HSI it is not possible to classify the scene to these five classes without ground truth data. Because of the limitation of HSI to maximum of three bands [33] that can be fused with another sensor data, it requires appropriate approach to select three optimal bands from hyperspectral data with dozens of bands. Another difference is that HSI product is not automatically classified as does DST, so it involves trial-and-error to identify best classification method with adequate training sites. Although DST offers better performance for urban feature extraction but its non-availability in the most commonly used image
processing software is a setback. HSI is very simple to implement and available in common image processing packages such as ENVI.

4. CONCLUSION

In this study, HSI and DST multi-sensor data fusion techniques were used to classify urban feature into five land cover types: metallic roof, non-metal roof, trees, grass, and paved surfaces. Quantitative and qualitative analysis shows that DST performs better than HSI method. Meanwhile developing DST as a module in commercial image processing packages will increase the effectiveness and accuracy of urban feature detection and extraction in the phase of today’s wide availability of remote sensing data from active and passive sensors. DST offers fast and cost effective solution for urban feature extraction and enhancement of the accuracy of classification without prior knowledge of the scene. With this, it is apparent that when optimal urban feature information extraction (including 3D) is desired, multi-sensor data fusion is the next thing to getting there.

5. REFERENCES