CHAPTER 4
DATA PREPROCESSING

4.1 DATA USED

The urban land use classification upto formal/informal settlements (refer Fig. 5.7) and their characterization with respect to spectral, contextual and structural properties were done using the IRS 1C PAN data and resolution merged data (IRS 1C PAN + LISS III). The fuzzy-driven knowledge base for urban land use classification was generated independently both for IRS 1C PAN data and IRS 1C LISS III +PAN merged data. The features of data used with respect to radiometric resolution and the procurement time are shown in detail in the Tables 4.1 and 4.2 separately. Figs. 4.1 and 4.2 show the LISS III and PAN data used for this study.

The PAN data was enhanced with image sharpening and LISS III data was merged with PAN. Two different methods namely HUE-INTENSITY-SATURATION (HIS) TRANSFORMATION and PRINCIPAL COMPONENT TRANSFORMATION (PC) were applied for merging which are described in Section 4.2 under the sub category ‘Data Fusion’.
Fig. 4.1 LISS III data of IRS 1C used for the study

Fig. 4.2 Sharpened PAN of IRS 1C used for the study
### Table 4.1 Details of digital data used

<table>
<thead>
<tr>
<th>Sl. No.</th>
<th>Data used</th>
<th>Year of pass</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>IRS 1C PAN</td>
<td>June 1998</td>
<td>NRSA Data Center, Hyderabad, Institute for Water Studies, Chennai And Space imaging Inc. USA</td>
</tr>
<tr>
<td>2</td>
<td>HIS merged</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>IRS 1C PAN &amp; LISS III</td>
<td>June 1998 &amp; May 1998</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>PC merged</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>IRS 1C PAN &amp; LISS III</td>
<td>May 1998 &amp; June 1998</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>IKONOS -2</td>
<td>January 2001</td>
<td></td>
</tr>
</tbody>
</table>

### Table 4.2 Sensor characteristics of digital data used

<table>
<thead>
<tr>
<th>Sl. No.</th>
<th>Sensor Character</th>
<th>Sensor characters of</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IRS 1C PAN</td>
<td>IRS 1C LISS III</td>
</tr>
<tr>
<td>1</td>
<td>Spatial Resolution (m)</td>
<td>5.8</td>
</tr>
<tr>
<td>2</td>
<td>Radiometric Resolution (bits)</td>
<td>6</td>
</tr>
<tr>
<td>3</td>
<td>Spectral Resolution (μm)</td>
<td>0.5 - 0.75</td>
</tr>
</tbody>
</table>
4.2 DATA PREPROCESSING

4.2.1 PAN Sharpening

The information content of the PAN image was improved by radiometric enhancement. For this, a quasi-uniform histogram was produced by flattening the non-uniform histogram (equation 4.1). For an image containing a total of N pixels and L brightness values, the equalized histogram should have N/L pixels associated with every brightness value called L. Therefore,

$$\frac{dy}{dx} = \frac{D(f(x))}{dx} = \frac{(L-i)}{N}$$

(4.1)

in which y = f(x) is the sought-for mapping or transformation if brightness values that take the original histogram of an image into an uniform histogram. The result is shown in Fig. 4.1.

4.2.2 Data Fusion

In multispectral data, the image enhancement was achieved by image fusion, otherwise called “resolution merging”, where spectral advantage of the multispectral LISS III and the spatial advantage of PAN data sets were exploited. The transformation techniques for resolution merging of the remotely sensed data were based on the change of the actual color space into another space and replacement of the one of the new gained component by a more highly resolved image. The basic principle of the resolution merging techniques is shown in Fig. 4.3. Two methods of resolution merging used in this study are discussed below.
4.2.2.1 Hue-Intensity-Saturation (HIS) Transformation

The **Hue-Intensity-Saturation (HIS) Transformation** system offers an advantage that the separate channels outline certain color properties, namely, Intensity (I), Hue (H) and Saturation (S). This specific color space is often preferred because the visual cognitive system of human being tends to treat three components as roughly orthogonal perceptual axes. The following transformation equation describes derivation of the HIS components. The transformation applied for this method is detailed in equation 3.2.

The intensity, I, describes the total color brightness and exhibits a strong similarity to the more highly resolved panchromatic image. Hence, it is
replaced with PAN data and the merged result is converted back to the RGB (XS₃, XS₂, XS₁) space.

\[
\begin{pmatrix}
I \\
V₁ \\
V₂
\end{pmatrix}
= \begin{pmatrix}
1 & 1 & 1 \\
3 & 3 & 3 \\
-1 & -1 & 2 \\
\sqrt{6} & \sqrt{6} & \sqrt{6} \\
1 & 1 & 0 \\
\sqrt{6} & \sqrt{6} & 0
\end{pmatrix}
\begin{pmatrix}
XS₃^R \\
XS₂^R \\
XS₁^R
\end{pmatrix}
\]

where \( H = \tan^{-1} \frac{V₂}{V₁} \) and

\( S = \sqrt{(V₂ + V₁)} \)

Fig. 4.4 explains pictorially the HIS merging method and Fig. 4.5 depicts the resolution merged data product using HIS transformation.

---

**Fig. 4.4 Hue-Intensity-Saturation (HIS) merging technique**
Fig. 4.5 Hue-Intensity-Saturation (HIS) resolution merged data (LISS III+PAN)
4.2.2.2 Principal Component (PC) Transformation

In Principle Component (PC) transformation for resolution merging, the multispectral bands are transformed and after replacing the first component by the high resolution data the fused imagery is back projected. The transformation is based on the statistical properties of the image using the covariance matrix $C_f$. The function $E$ in equation 3.3 refers to the expectation that:

$$C_f = E((f - \mu_f) \cdot (f - \mu_f)^T)$$  \hspace{1cm} (3.3)

where the variables $f$ and $\mu_f$ are the vectors formed by the lexicographically ordered image and the mean images, respectively. The ordered eigenvectors (with respect to value of the eigenvalues) of $C_f$ generate the transformation matrix. Multiplication with the image vector $f$ transforms the multispectral imagery into a space that minimizes the correlation between the bands and allows an optimal replacement. The technique is illustrated in Fig. 4.6 and the result in Fig. 4.7.

Fig. 4.6 Principal Component (PC) merging technique
Fig. 4.7 Principal Component (PC) resolution merged data (LISS III+PAN)
Preprocessing involved in the *data preparation* of IRS 1C PAN and LISS III data for object generation and classification is shown in Fig. 4.8.

**Fig. 4.8 Data preprocessing for object generation**

### 4.2.3 IKONOS Data

PAN sharpened IKONOS data combining the spatial detail of the 1 meter panchromatic sensor with the colour information of 4 meter multispectral sensor to produce a 3-band multispectral image at 1 meter sampling distance was used for this study. Natural colour orders were filled with red, green and blue in their corresponding TIFF positions. The colour enhanced IKONOS data covering a part of Anna University and its surroundings used for object based image analysis is shown in Fig. 4.9.

An object based classification using objects and units instead of pixels was attempted on the sharpened IRS PAN, HIS-merged IRS 1C data and PAN-sharpened IKONOS data. In the present study, knowledge base representing the spectral, spatial and contextual keys for classification of individual urban classes (refer Section 5.1) was constructed for this purpose. The classification achieved
Fig. 4.9. PAN sharpened IKONOS data used for the study.
was verified for thematic accuracy and the classification accuracy to evaluate object based methodology to ascertain its applicability to various sensor and data. The evaluation for object based classification method was done with reference to urban land use classification, as this is most heterogeneous and dynamic land use phenomenon.