

## LESSON 9 – Address Geocoding

In this class we will show how Terra View can be used in the Address Geocoding process. This tool can be used whenever point data need to be considered from the spatial perspective but its geographical coordinates are not available and only the addresses are known. The objective of address geocoding is to establish, by estimating it, the geographical position corresponding to each address. After this process, the points are said to be spatialized, and can be analysed or just visualized on maps.

A set of data referring to the locations in the city of Taubaté, in the state of São Paulo will be used in this lesson. Before starting, please import the data from the Taubate\_SP.shp file in the “Dados do Tutorial” folder. Such data are given in LatLon projection with datum SAD69.

### 9.1 How Address Geocoding works

It is absolutely essential to the address geocoding process the availability of a set of spatial data which consists basically of an information layer whose objects are location segments, each one containing the address ranges (numbers) of one or both sides of the segment. This layer will be used as a **reference** or **basis** for the geocoding process. Therefore, the method for finding the location from an address consists basically of:

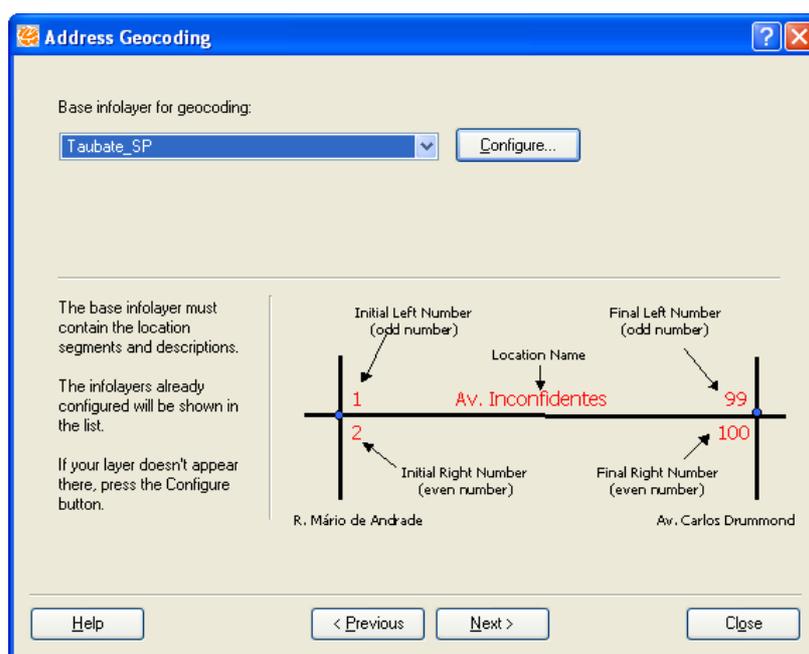
1. Finding the location segment whose name is equal or closest to the name shown in the address and having one side with a number range that encompasses the address number.
2. By means of linear interpolation, to calculate the geographical position based on the length of the segment found in the previous step, on its number range and on the address number.

In several situations this method may not result in locating the geographical position such as when no segment has been found (for instance, due to a spelling error) or when several segments have been found (for instance, due to overlapping of their number ranges, all of them encompassing the address number).

## 9.2 The Address Geocoding wizard

Most of the Address Geocoding process is accomplished in TerraView by using a wizard. This makes possible to configurate one or more reference plans for address geocoding (see steps 1-3 below), and then to import - in case there isn't one yet -, a table called *geocoding table* that has the addresses to be geocoded and to accomplish the geocoding of its addresses (steps 4-17)

In order to start this wizard use the **File** → **Address Geocoding...** option and the window below (Figure 9.1) will be shown.



**Figure 9.1 – Starting the Geocoding wizard.**

Follow the steps below:

1. Select the information layer to be used as reference in the geocoding process, that is, the layer containing the location segments, as explained in item 9.1. For geocoding purposes, only the layers that have been previously configured as reference will be shown in the list. If the desired layer has already been configured, select it from the list and click **Next** button, which will be enabled. You will be taken to step (4) below. If the desired layer has not been configured yet, click **Configure** button, then proceed to step (2).
2. As shown in **Configure Layer as Reference** window (Figure 9.2), it is possible to configure an information layer as reference in the geocoding process. The configuration consists of defining the fields that will be used for the location segments of the information layer. Proceed as detailed below:

- a) One or more fields associated to the location name must be informed by selection. This name can be divided into a maximum of four fields corresponding to the type, title, preposition and name of the location, if available.
- b) A field with the full name of the location is required in order to assure a good geocoding process performance. If it already exists, it should be selected in the left field of the **Select Full Name** line, otherwise it should be created, just by writing its name in the field to the right of such line. In such case, the field will be created and receive the contents of all fields that compose the location name.
- c) If available, the fields corresponding to the location segment numbers such as initial odd number, initial even number, final odd number and final even numbers must be informed next.
- d) Filling out the other fields of this window is optional. After completing the information, click the **Confirm** button.

From this moment on, the information layer can be used as reference in the address geocoding process, and will be shown in the window of Figure 9.1. In other words, you will be sent back to step (1).

**Configure layer as reference for geocoding**

Infolayer:  
 Taubate\_SP

Reference Fields for Address Location

Location type: TIPO      Location title: [Empty]

Location preposition name: [Empty]      Location Name: NOME

Select the complete name or enter with a new column name:  
 NOME\_COMPLETO      [Empty]

Initial odd address: NUM\_IE      Initial even address: NUM\_ID

Final odd address: NUM\_FE      Final even address: NUM\_FD

Location left neighborhood: BAIRRO\_LE      Location right neighborhood: BAIRRO\_LD

Location left zipcode: CEP\_LE      Location right zipcode: CEP\_LD

Buttons: Help, Confirm, Remove, Close

**Figure 9.2 – Configuring layer as reference.**

3. When a layer that has already been configured as reference for the geocoding process is selected in the **Configure Layer as Reference** window (Figure 9.2), it will make possible to do the following: a) change any of the previously selected fields, allowing for the information to be updated by clicking the **Confirm** button; b) remove the layer from the geocoding reference layers list, by clicking the **Remove** button. Once the information layer has been configured it will no longer be necessary to repeat these steps on future address locations, including after closing the locate window (that will be seen on Figure 9.12) or TerraView.
4. To come to this step, you have clicked the **Next** button as shown in the Figure 9.1 window. Here just a window explaining the next step – that consists of presenting a table with addresses to be geocoded, will be shown. You will get to the Figure 9.3 window upon clicking the **Next** button.

**Figure 9.3 – Details of the table to be geocoded.**

5. The window in Figure 9.3 shows details of the table containing the set of addresses to be geocoded. This table should have been previously imported to the database as being of the *geocoding data* type, one of the valid types in the interface for table importing, as shown in a previous lesson. The geocoding data tables that have already been imported can be picked up from the **Table to geocode** list.

NOTE: In case the desired table has not been imported yet, you can actuate the respective interface by clicking the **Import Table** button. Do this in order to import the ENDERECO.DBF file from the “Dados do Tutorial” folder. For any doubts, please refer to item 4.2.1 of lesson 4. The only difference is that it is already presumed it’s a table of the *geocoding data* type, therefore it is not necessary to mention the type of table.

6. Provide the details on the geocoding table as shown in Figure 9.3 and click the **Next** button.
7. In this step a dictionary can be used during the address search on the reference information layer, in order to facilitate the search process. The dictionary must have words in the reference information layer that are written differently on the addresses to be geocoded (see Figure 9.4). Please refer to the example below to learn how the dictionary works and how it is structured.

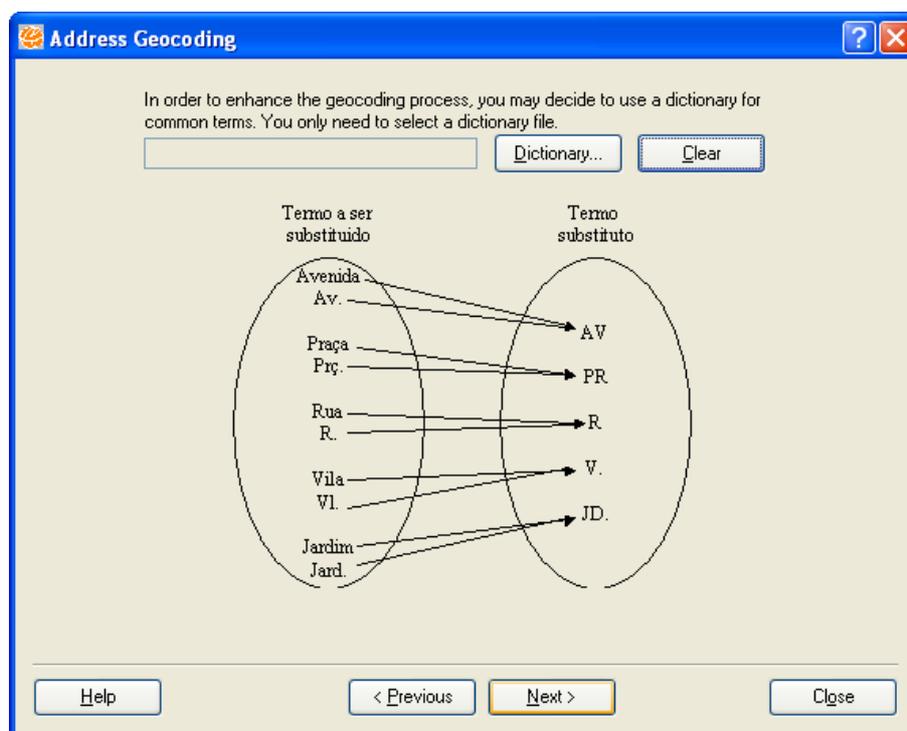
Example: The address list could show the following location name: RUA CÔNEGO CASSIANO and on the reference information layer, R CON. CASSIANO. Thus, if a line in the dictionary linking the word RUA to the letter R and another linking CÔNEGO to CON. exists, the search can be successful. In this case and according to Figure 9.4, it could have the following text lines:

”AV”, ”AV.”, ”AVENIDA”

”CON.”, ”CÔNEGO”

”R”, ”R.”, ”RUA”

In this format the first word will be used when others appear in the same line, too. In the above example, both ”R.” and ”RUA” can be replaced by ”R” in the search (however, this will only happen after they are not found the first time).



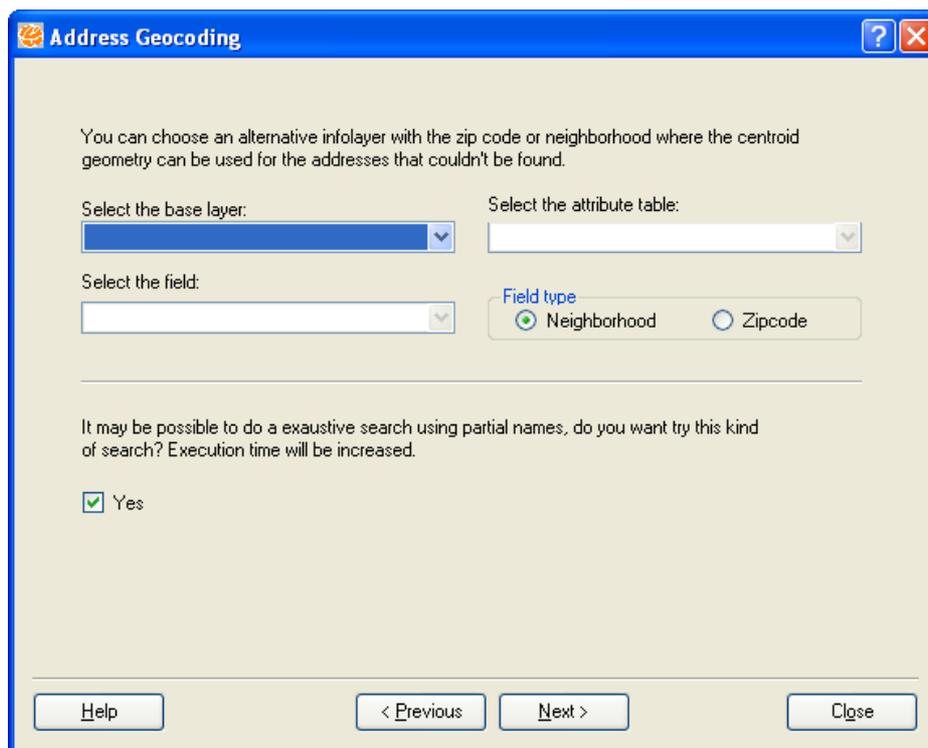
**Figure 9.4 – Selecting the dictionary file.**

8. Click on the **Next** button
9. The next window allows the replacement of special characters in the input (addresses to be decoded) to be used when the reference information layer does not have such characters.



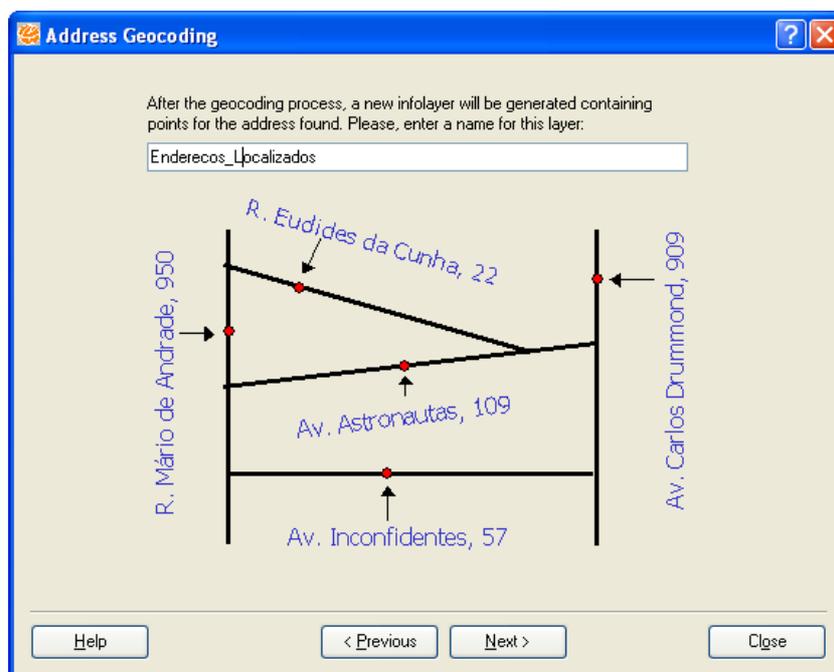
**Figure 9.5 – Replacing special characters.**

10. Click on the **Next** button.
11. In the Figure 9.6 window, you may choose an information layer with objects containing centroids that will be used for determining the geocoded points if the search by address has not been successful. It should be informed whether the geocoding modules in this window will exhaustively search the basis – a procedure that although making the geocoding process longer, improves the level of address finding (for example, by similarities in names). Select this choice.



**Figure 9.6 – Centroid use and exhaustive search.**

12. Click on the **Next** button.
13. In order to finish the address geocoding process, it is required that the name of the new information layer to be produced from the geocoding process be written (refer to Figure 9.7).



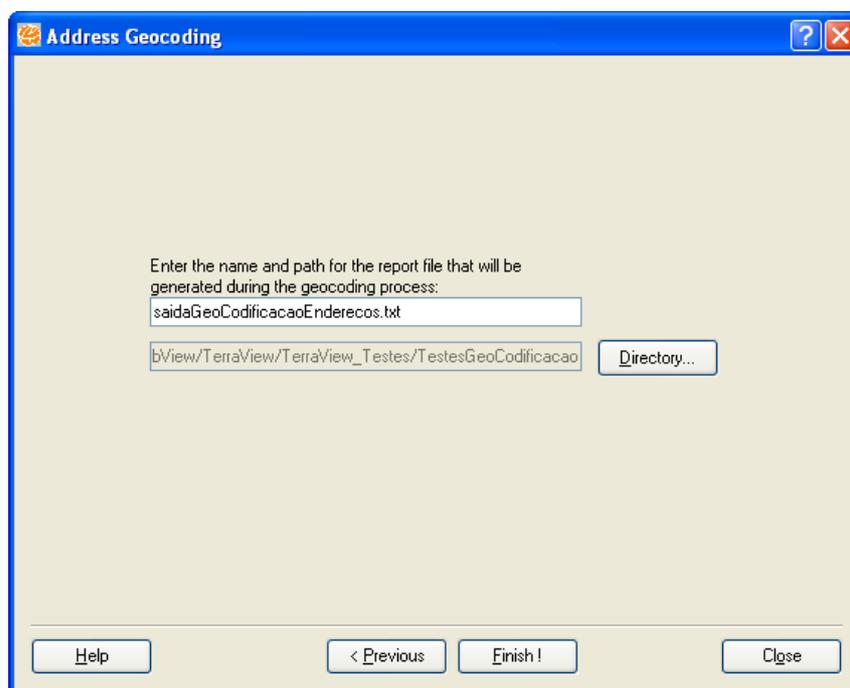
**Figure 9.7 – Name of the new layer with geocoded points.**

National Institute for Space Research - INPE – Image Processing Division



<http://creativecommons.org/licenses/by-nc-sa/2.5/deed.pt>  
Esta obra está licenciada sob uma Licença Creative Commons

14. Click on the **Next** button.
15. The name of the directory and file that will show the details in carrying out the geocoding process (Figure 9.8) should be entered next.



**Figure 9.8 – Geocoding Process Report file.**

16. For each address searched, the report shows the code resulting code from the geocoding process, the line identifier of each address searched and, in case it has been found, the attributes of street segment whose range encompasses the number of the address searched. The existing codes and what they mean are shown below:

- ✓ 0: Database connection problem
- ✓ 1: Address not found
- ✓ 2: Empty full name
- ✓ 3: Database query error
- ✓ 4: Address found by name and number
- ✓ 5: Multiple addresses found by name and number
- ✓ 6: Address found by name and no number
- ✓ 7: Multiple addresses found by name and no number
- ✓ 8: Address found by similarity
- ✓ 9: Address found by neighborhood or postal code
- ✓ 10: Address found by name, number, and no type
- ✓ 11: Multiple addresses found by name, number, and no type
- ✓ 12: Address found by name, no number, and no type
- ✓ 13: Multiple addresses found by name, no number, and no type
- ✓ 14: Address found by name, number, no type, and no title
- ✓ 15: Multiple addresses found by name, number, no type, and no title

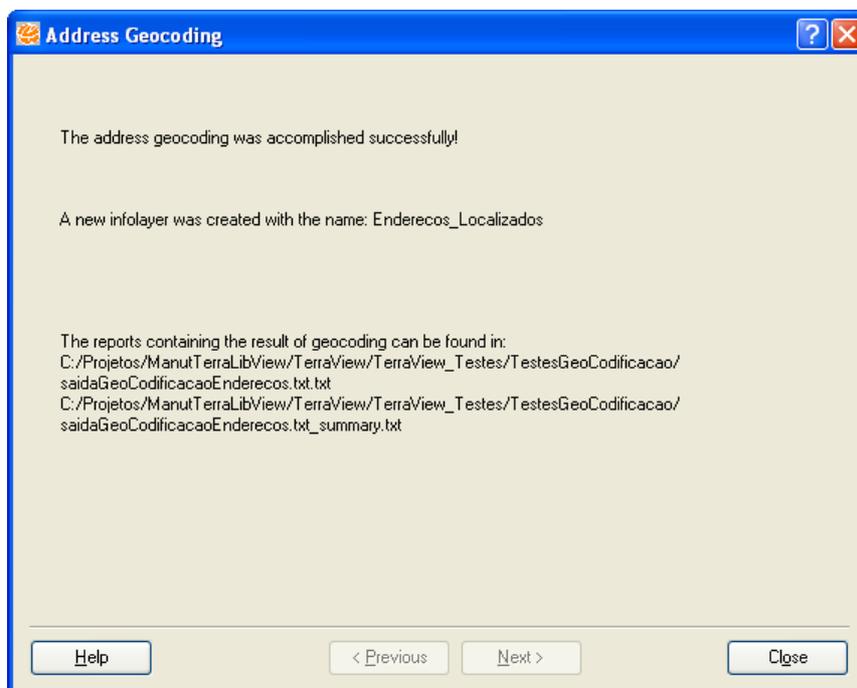
- ✓ 16: Address found by name, no number, no type, and no title
- ✓ 17: Multiple addresses found by name, no number, no type, and no title
- ✓ 18: Address found by name, number, no type, no title, and no preposition
- ✓ 19: Multiple addresses found by name, number, no type, no title, and no preposition
- ✓ 20: Address found by name, no number, no type, no title, and no preposition
- ✓ 21: Multiple addresses found by name, no number, no type, no title, and no preposition

NOTE 1: Codes greater than 10 apply only in exhaustive search.

NOTE 2: In addition to a detailed report, a summary with the total number of addresses searched corresponding to each of the codes listed above, and the overall totals of addresses searched and found is also produced.

17. Click on the **Finish** button.

18. The window to close the address geocoding process will appear on the screen (Figure 9.9). Click on **Close** button.



**Figure 9.9 – Closing the geocoding wizard.**

19. Once the process is finished, the layer with points corresponding to the addresses which have been geocoded can be visualized, as shown in Figure 9.10.

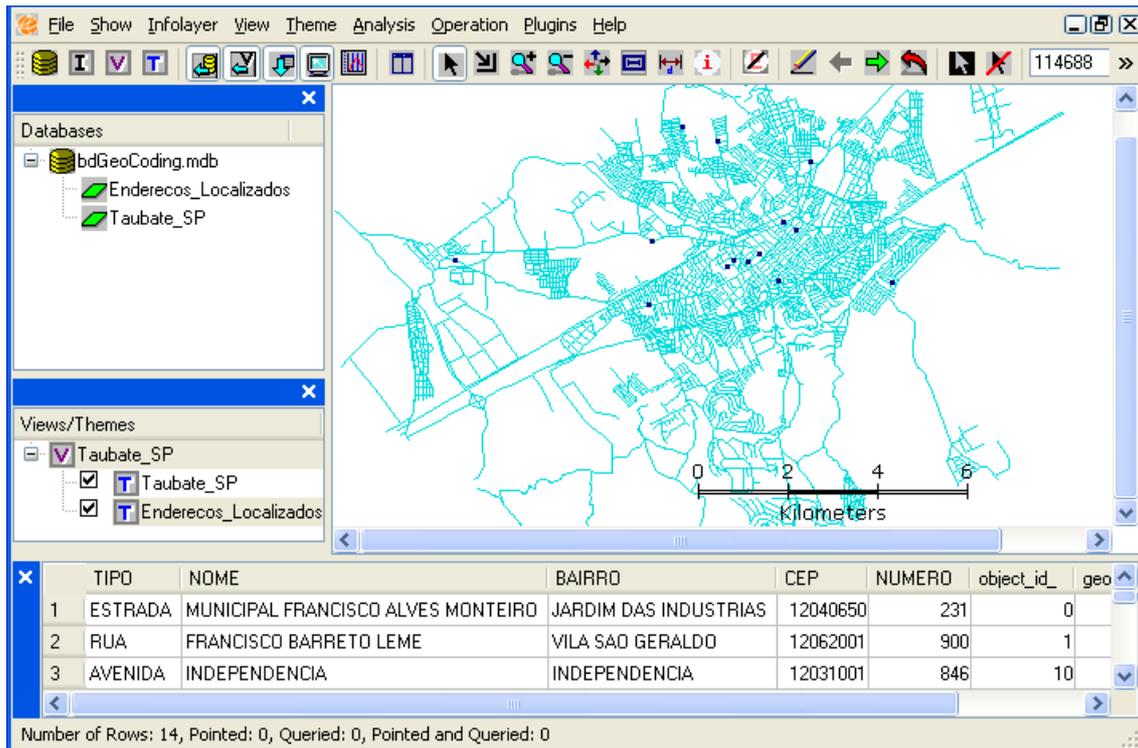
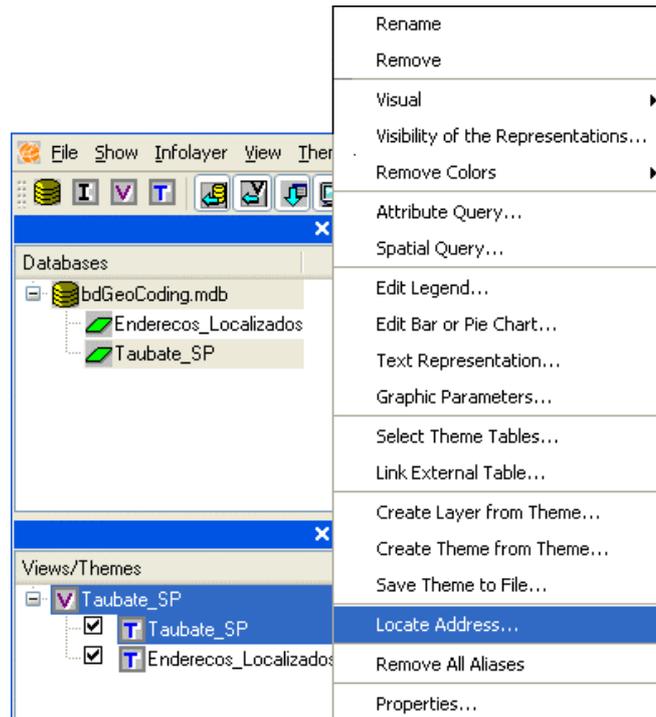


Figure 9.10 – Resulting layer showing georeferenced addresses.

### 9.3 How to Locate an Address

This feature of address geocoding is not provided by the Address Geocoding wizard studied in the previous item. It makes possible for an address to be found and be seen on a temporary basis on the visualization area. Follow these steps:

1. Select a theme corresponding to the address geocoding reference layer that has been configured and used in the previous item (Taubate\_SP). In case there is no such theme, one should be created so as to make use of this feature. Open the context menu for this theme, by clicking the right button of the mouse (Figure 9.11).



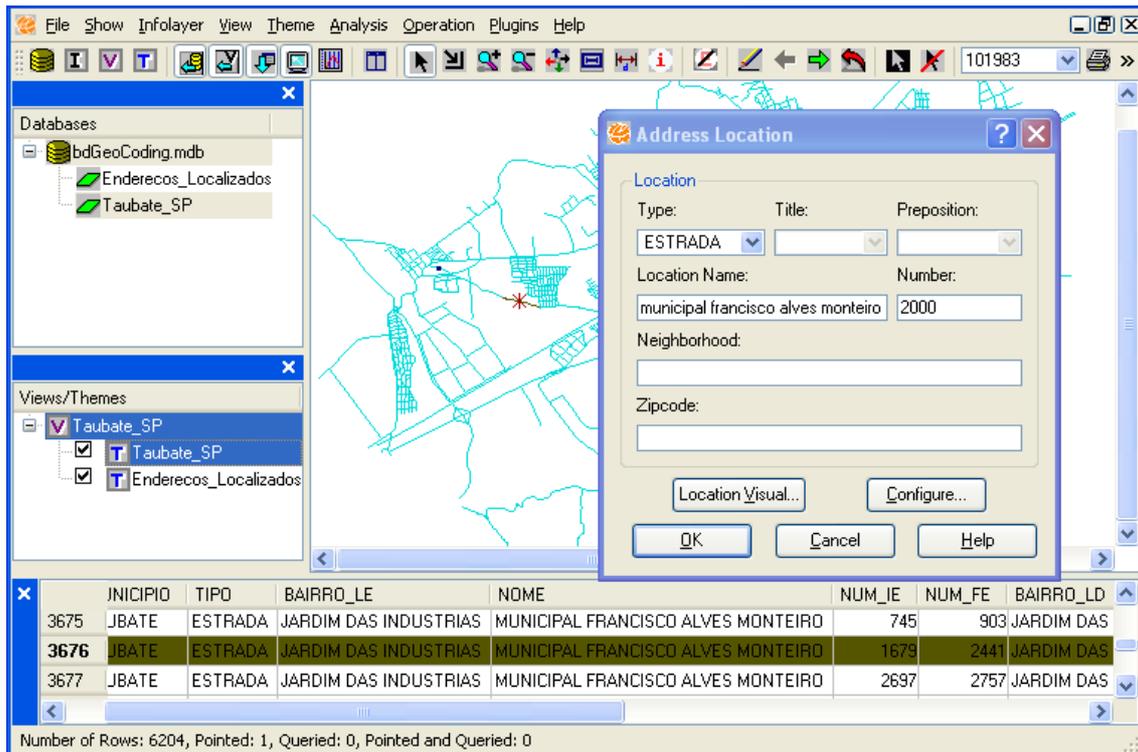
**Figure 9.11 – Activating the address location interface.**

2. Type the type, title, preposition and name of the location as well as the address number.

**Figure 9.12 – Address Location interface.**

The **Configure** button on the address location window can be selected. This will open a window from the previous section, so that the configuration of the information layer of the theme as a reference layer for geocoding can be adjusted.

3. Click on the **OK** button.
4. The requested address will be shown as a dot on the screen (Figure 9.13). The shape, color or size of this dot can be changed by clicking the **Location Visual...** button. In the example given, the dot was changed to a star shape.



**Figure 9.13 – Address found.**