Model-Based Recognition of Domino Tiles and other Objects using TGraphs

Stefan Wirtz

wirtzstefan@uni-koblenz.de

Institute for Computational Visualistics
University of Koblenz and Landau

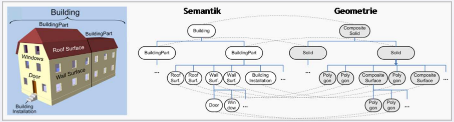
17. Juli 2009

Overview

Model-based object recognition in 2D images

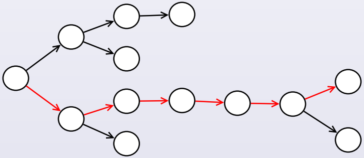
Given:

- 2D images in a scene
- Model as graph:
  - Declarative knowledge
  - Procedural knowledge



Searched:

- Recognition of the model in the image
  - Application independent
  - Shortest path in the searchspace
  - Rating



Application domains are:

- Domino tiles
- Poker cards
- Campus respectively buildings

Motivation I

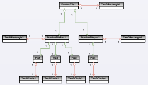
- Object recognition using model-based image analysis, based on symbolic descriptions, is still a challenge.
- Drawback is that many knowledge-based models are needed.
- Google Earth provides lots of user generated models of buildings which are designed in Geography Markup Language (GML) and in Keyhole Markup Language (KML).

Motivation II

- Research in the model-based Image analysis is meaningful.

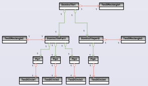
Given:

- Image(s).
- Set of models.

, , , . . . , 

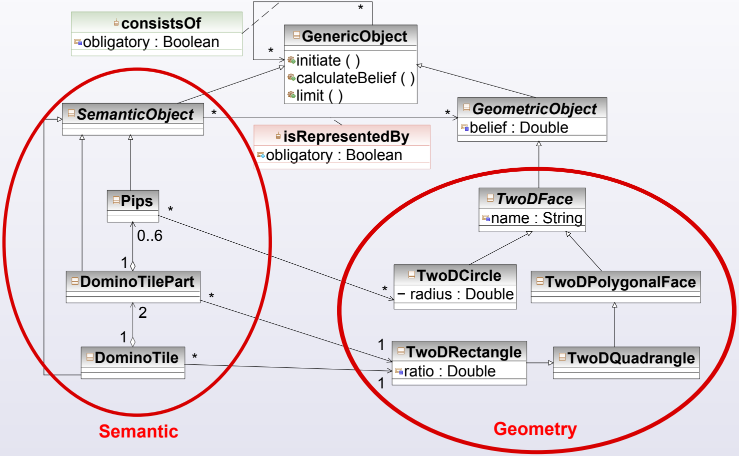
Searched:

- Rating of the models and assignments.
- Set of possible models.

, , . . . , 

$p = 0.1$ $p = 0.7$ $p = 0.2$

Model scheme



- Separation of semantic and geometry

Knowledge

In addition to declarative knowledge, we also model the procedural knowledge in the scheme:

- instantiate()
- limit()
- calculateBelief()

With this scheme, we are able to construct concrete models. In this case study, we have 28 basic models, because we have 28 domino tiles.

TGraph

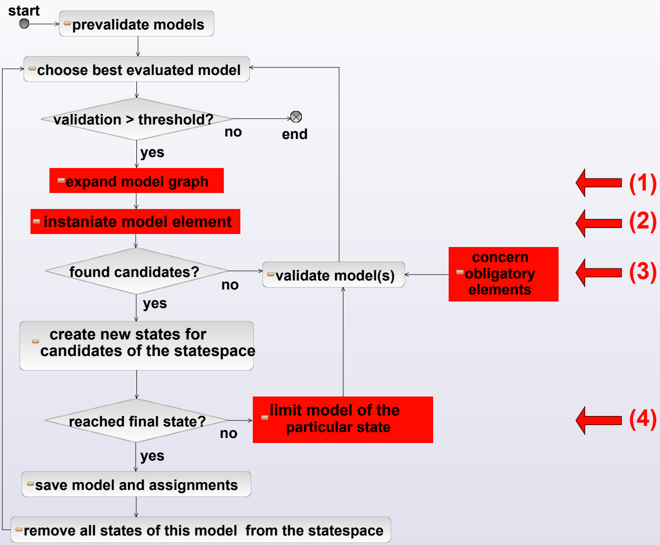
- TGraphs [ERW08] are used as a “light-weight”-ontology to construct model schemes.
- TGraphs are a very generic graph concept based on vertices and edges as first-class entities.
- TGraphs includes types, attributes and ordering for both.

ERNEST

One archetype of such a system is ERNEST [NSSK90]

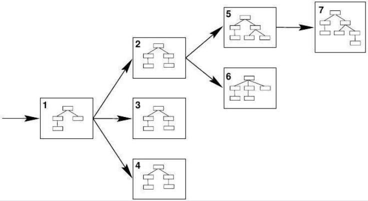
- A pattern recognition system developed in the eighties.
- Employing semantic networks in several application domains:
  - Building recognition [Qui97]
  - Speech understanding [Nie90]
- In ERNEST six task-independent rules precisely define the computation of elements and instances.
- These rules are combined with graph search algorithms to handle the control problem.

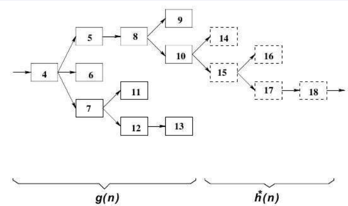
System


start
prevalidate models
choose best evaluated model
validation > threshold?
no
end
yes
expand model graph
instaniate model element
found candidates?
no
validate model(s)
concern obligatory elements
yes
create new states for candidates of the statespace
reached final state?
no
limit model of the particular state
yes
save model and assignments
remove all states of this model from the statespace
(1)
(2)
(3)
(4)

A*-Algorithm

- Controls the application flow in the state space.
- $f(n) = g(n) + \epsilon\, h(n)$
- $g(n)$ is the belief of the already assigned elements.
- $h(n)$ is the estimated belief of the not yet assigned elements.
- $h(n)$ has to be better than the real belief, therefore we choose as estimation perfect assignments.





Dempster Shafer

- Dempster-Shafer provides also a convenient combination rule for independent beliefs.
- A *basic probability assignment* (τ) is needed and τ must fulfill the conditions:
  - Given the sample space $\Omega = \{a_k\}$,
  - $k = 1, \ldots, K$, $A \subseteq \Omega$, where a_k and A are arbitrary events.

$$\tau(\emptyset) = 0 \qquad (1) \qquad\qquad \sum_{A \subseteq \Omega} \tau(A) = 1 \qquad (2)$$

Dempster Shafer I

The combined τ with Dempster-Shafer is:

$$\tau_1 \oplus \tau_2 = \tau(A) = \begin{cases} 0 & : \quad A = \emptyset \\ \frac{\sum_{t \cap u = A} \tau_1(t)\tau_2(u)}{1 - \sum_{t \cap u \neq \emptyset} \tau_1(t)\tau_2(u)} & : \quad \emptyset \neq A \subseteq \Omega \end{cases} . \tag{3}$$

- The Dempster-Shafer rule is commutative and associative.
- The resources have to be independent of each other because in general $\tau \oplus \tau \neq \tau$ is valid.

General Functions I

Therefore, the combination rule of DS is used like Quint made it in [Qui97]. Then the belief of a `SemanticObject` is:

$$\tau(\texttt{SemanticObject}|\mathrm{E}_{parts}, \mathrm{E}_{rep}) =$$
$$\kappa_1\ \tau(\texttt{SemanticObject}|\mathrm{E}_{parts}) \oplus \kappa_2\ \tau(\texttt{SemanticObject}|\mathrm{E}_{rep})\ .$$

- With $\kappa_1, \kappa_2 \in [0, 1]$ you can weight the trust that this event really support the belief in the `SemanticObject`.
- The information source E_{parts} describes the observation of the parts of a `SemanticObject`.
- The information source E_{rep} specifies the GeometricObject which represents this `SemanticObject`.

General Functions I

Therefore, the combination rule of DS is used like Quint made it in [Qui97]. Then the belief of a `SemanticObject` is:

$$\tau(\texttt{SemanticObject}|\mathrm{E}_{parts}, \mathrm{E}_{rep}) =$$
$$\kappa_1 \, \tau(\texttt{SemanticObject}|\mathrm{E}_{parts}) \oplus \kappa_2 \, \tau(\texttt{SemanticObject}|\mathrm{E}_{rep}) \, .$$

- With $\kappa_1, \kappa_2 \in [0, 1]$ you can weight the trust that this event really support the belief in the `SemanticObject`.
- The information source E_{parts} describes the observation of the parts of a `SemanticObject`.
- The information source E_{rep} specifies the GeometricObject which represents this `SemanticObject`.

General Functions II

We use a function equal to the total-probability rule.

$$\tau(\texttt{SemanticObject}|\mathrm{E}_{\mathrm{parts}}) = \sum_j \tau(\texttt{SemanticObject}|\mathrm{B}_j)\tau(\mathrm{B}_j|\mathrm{E}_{\mathrm{parts}}) \,.$$

For any partition B_j, $j = 1, \ldots, \mathrm{N}$ of the event space Ω and $\Omega = \{\{\mathrm{fp}_1, \ldots, \mathrm{fp}_\mathrm{L}\}, \{\overline{\mathrm{fp}_1}, \ldots, \mathrm{fp}_\mathrm{L}\}, \ldots, \{\overline{\mathrm{fp}_1}, \ldots, \overline{\mathrm{fp}_\mathrm{L}}\}\}$, where fp equates the event of found a `SemanticObject` part.

General Functions III

- Finally, we need the belief of a partition B_j given the information source $\mathrm{E}_{\mathrm{B}_j}$:

$$\tau(\mathrm{B}_j|\mathrm{E}_{\mathrm{B}_j}) = \prod_{p_j \in \mathrm{B}_j \wedge p_j \mathrm{isTrue}} \tau(p_j) \; .$$

Specific Functions I

- Specific functions for the assignments have an big impact on the classification.

For example:

- The belief function which penalize missing assignments of segmentation objects and/or model elements:

$$\tau(\mathrm{dr}_i|\mathrm{P}) = e^{-\frac{1}{2}x^2}$$
$$x = ||\mathrm{C}| + |\{\text{given pips}\} \in \mathrm{P}| - |\{\text{associated pips}\} \in \mathrm{P}|| \ .$$

Specific Functions II

- The belief in the assignment of a pip p to a circle c:

$$\tau(p \in \mathrm{P} | c \in \mathrm{C}) = 101 - 200^{q} - 50|c_r - p_r|,$$
$$q = \sqrt{(p_x - c_x)^2 + (p_y - c_y)^2}\ .$$

- Where (c_x, c_y) and (p_x, p_y) are the center points of the circles, and c_r and p_r are the radiuses, is very important.

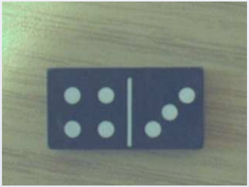
Data set

- The data set containing 489 images of single domino tiles on a homogeneous background:

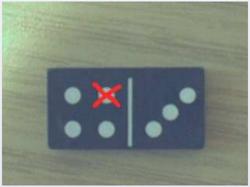
Results

Classification rates

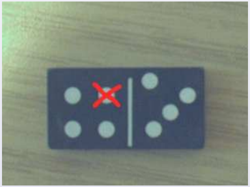
detection	precision	recall
94.7%	94.4%	89.4%



4-3 domino tile
$\tau = 97.3\%$



4-3 domino tile
$\tau = 84.9\%$



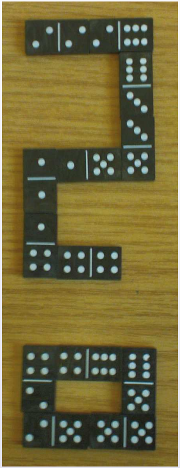
4-4 domino tile
$\tau = 68.9\%$

Conclusion

- We can calculate the belief in all assignments.
- Also we obtain the knowledge which elements of an image is the best correspondence for an element among the model.
- To adapt this strategy to building recognition we have to create 3-D models.
  - Out of this 3D models, according to estimated perspectives, 2D models will be created.

Ende

“It is proven that the celebration of birthdays is healthy. Statistics show that those people who celebrate the most birthdays become the oldest.”S. den Hartog, Ph D. Thesis, Universtity of Groningen.



EBERT, Jürgen ; RIEDIGER, Volker ; WINTER, Andreas:
Graph Technology in Reverse Engineering, The TGraph Approach.
In: GIMNICH, Rainer (Hrsg.) ; KAISER, Uwe (Hrsg.) ; QUANTE, Jochen (Hrsg.) ; WINTER, Andreas (Hrsg.): *10th Workshop Software Reengineering (WSR 2008)* Bd. 126. Bonn : GI, 2008, S. 67–81

NIEMANN, Heinrich:
Pattern Analysis and Understanding. Bd. 4.
Heidelberg : Springer Verlag, 1990

NIEMANN, H. ; SAGERER, G. ; SCHRÖDER, S. ; KUMMERT, F.:
ERNEST: A Semantic Network System for Pattern Understanding.
In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 9 (1990), S. 883–905

QUINT, Franz:
Kartengestützte Interpretation monokularer Luftbilder.
Karlsruhe, Universität Fridericiana zu Karlsruhe (TH), Diss., 1997

Discussion

Are there any ideas how to handle such problems?

Task specification

Given:

- a universe $\mathrm{SEGMENT}$ of all possible segmentation objects,
- a set $\mathrm{RECT} \subset \mathrm{SEGMENT}$ of all possible rectangles,
- a set $\mathrm{CIRCLE} \subset \mathrm{SEGMENT}$ of all possible circles,
- $\mathrm{CIRCLE} \cap \mathrm{RECT} = \emptyset$
- a relation $\mathrm{CONTAINS}$ which describes if a segmentation objects contains another segmentation objects, $\mathrm{CONTAINS} \subseteq \mathrm{SEGMENT} \times \mathrm{SEGMENT}$,

Given:

- a set $\mathrm{P} \subset \mathrm{CIRCLE}$ of possible pips with a specific position, with $|\mathrm{P}| = 14$,
- a set LAYOUT of possible layouts, $|\mathrm{LAYOUT}| = 28$, the number of possible domino tiles, and $\mathrm{LAYOUT} \subset 2^{\mathrm{P}}$,
- a set of Models $\mathrm{M} = \{\mathrm{dr}_1\} \times \{\mathrm{dr}_2\} \times \{\mathrm{dr}_3\} \times \mathrm{LAYOUT}$, where $m \in \mathrm{M} \wedge m = (m_1, m_2, m_3, m_4)$ is a concrete model.

Input:

- a set $\mathrm{S} \subset \mathrm{SEGMENT}$, $\mathrm{S} = \mathrm{R} \dot{\cup} \mathrm{C}$ with $\mathrm{R} \subset \mathrm{RECT}$ and $\mathrm{C} \subset \mathrm{CIRCLE}$, of all segmentation objects found in the image,
- a threshold $\theta \in [0, 1]$, to set a belief limit of the hypotheses
- a total function $\mathrm{Bel}_{\mathrm{S}} : \mathrm{S} \to [0, 1] \subset \mathbb{R}$, describing the belief in the segmentation object itself
- a total function $\mathrm{Bel}_{\mathrm{M}} : \mathrm{M} \times (2^{\mathrm{R}} \times 2^{\mathrm{C}}) \to [0, 1] \subset \mathbb{R}$, describing the belief in the model assignment.

Output:

- a set $\mathrm{M}' \subset \mathrm{M}$ of domino models, where $\forall m' \in \mathrm{M}' | \mathrm{Bel}_{\mathrm{M}}(m', \mathrm{R}', \mathrm{C}') \geq \theta$, with $\mathrm{C}' \in 2^{\mathrm{C}} \wedge \mathrm{R}' \in 2^{\mathrm{R}}$,
- $\forall m' \in \mathrm{M}'$
  - a set $\mathrm{R}^{m'} \subset \mathrm{R}$ of rectangles partially mapped to m'_1, m'_2 and m'_3, with $|\mathrm{R}| < 4$, where $r_1 \in \mathrm{R}^{m'}$ is assigned to the domino tile m'_1,
  - $\gamma_{\mathrm{C}}^{m'} = \mathrm{C} \rightarrowtail\!\!\!\!\!\!\rightarrow m'_4$,
  - $\forall c_1 \in \mathrm{C} | c_1 \in \mathrm{dom}(\gamma_{\mathrm{C}}^{m'}) : \ r_1 \,\mathrm{CONTAINS}\, c_1$
  - judgments $\mathrm{J}_{\mathrm{local}_k}^{m'} \in [0, 1], \ k = 1, \ldots, \mathrm{L}$ how well these mappings in detail could be established,
  - a total judgment $\mathrm{J}^{m'} \in [0, 1]$ of the found models. These are calculated with $\mathrm{Bel}_{\mathrm{M}}$.

In this case study the judgment J is according to the belief of the Dempster-Shafer theory.