

## Análise de Imagens Baseada no Conhecimento: Métodos e Resultados

Prof. Dr.-Ing. Hans-Peter Bähr

Instituto de Fotogrametria e Sensoriamento Remoto – IPF  
 Universidade de Karlsruhe  
 Englerstr. 7, D-76128, Karlsruhe - Alemanha  
 ✉ baehr@ipf.bau-verm.uni-karlsruhe.de

<b>Conteúdo</b>	<ol style="list-style-type: none"> <li>1. Introdução: novas ferramentas conduzem a novos produtos</li> <li>2. O Conceito de Conhecimento</li> <li>3. Estado Atual da Análise de Dados com Base no Conhecimento, Mostrada Através dos Trabalhos Realizados no IPF/Universidade de Karlsruhe       <ol style="list-style-type: none"> <li>3.1 Classificação de Objetos em Imagens de Satélite de Alta Resolução (Segl, 1996)</li> <li>3.2 Classificação do Uso da Terra a partir de Imagens de Satélite e com Suporte de Cartas (K.-J. Schilling et al., 1996)</li> <li>3.3 Extração de Objetos em Fotos Aéreas com o Apoio de Cartas (F. Quint, 1997)</li> <li>3.4 Extração de Objetos em Cartas de Grande Escala (R. Brügelmann, 1998)</li> </ol> </li> <li>4. A Importância Substancial do Contexto</li> <li>5. Bibliografia</li> <li>6. Agradecimento</li> </ol>
-----------------	--

**Resumo:** *Sistemas de Informações Geográficas incluem bases de dados e modelos para a análise destes dados. Estes dois campos representam diferentes tipos de conhecimentos, quais sejam factuais e inferenciais ("fatos e regras"). Encontramos semelhante discriminação em muitas diferentes disciplinas.*

*Enquanto o conhecimento factual já tem sido amplamente descrito em termos de sua natureza determinística e estocástica, por sua vez é um grande desafio intelectual modelar-se o conhecimento inferencial ("inteligência artificial"). Esta publicação apresenta alguns resultados obtidos nestas áreas no Instituto de Fotogrametria e Sensoriamento Remoto da Universidade de Karlsruhe (IPF) ao longo dos últimos dois anos. As técnicas aplicadas variam desde redes neurais até a triangulação de Delaunay, redes semânticas e campos randômicos de Markoff e são bastante diversos. Gráficos são ideais para modelar o conhecimento, estando fatos representados nos nós e regras representadas pelas conexões.*

*O significado de uma característica de uma imagem está em seu contexto. Com respeito a isso, linguistas têm declarado que "o significado de uma palavra é o seu uso no idioma" (Wittgenstein, 1954). Portanto, entender uma imagem requer a descrição dos seus objetos, conjuntamente com a sua vizinhança.*

**Abstract:** *Geo-Information Systems include both data basis and models for data analysis. These two fields represent different kinds of knowledge, that is factual and inferential knowledge ("facts and rules"). We find similar discrimination in may different disciplines.*

*While factual knowledge has already been widely described for its deterministic and stochastic nature, it turns out as a big intellectual challenge to model inferential knowledge ("artificial intelligence"). The publication gives some results obtained in this field at the Institute for Photogrammetry and Remote Sensing of the Karlsruhe University (IPF) during the last two years. The employed approaches range from neural networks to Delaunay triangulation, semantic nets and Markoff random fields and seem to be very diverse. However it turns out, that in all cases the basic architecture is built by networks: Graphs are ideally suited for modeling knowledge, as facts are in the nodes and rules in the connections.*

*The meaning of a feature in an image is in its context. Accordingly, linguists had declared "the meaning of a word is its use in the language" (Wittgenstein, 1954). Therefore, understanding an image requires the description of its objects together with their neighborhood.*

### 1. Introdução: novas ferramentas conduzem a novos produtos

A aquisição sistemática e apresentação cartográfica de nosso meio-ambiente vêm sendo desenvolvidas já há muitos séculos. A aparência do produto mais importante, a carta, será diretamente influenciada pela ferramenta utilizada. Xilografia, litografia e impressão offset conduziam ao seus respectivos inconfundíveis produtos.

Através da aplicação de processamento digital defrontamo-nos atualmente com uma situação completamente diferente que, pela extensão das mudanças causadas, já foi classificada como um "Salto de Paradigma" (Ackermann, 1995). O passo dado agora entre a aquisição de dados por procedimentos geodésicos clássicos, com a apresentação dos dados em uma carta topográfica vetorial, e a aquisição digital de dados, reconhecimento automático de objetos e armazenamento em bancos de dados é muito maior do que aquele dado anteriormente, entre a mesa topográfica e estações totais para medidas no terreno.

O desafio das novas ferramentas é procurar restituir dados de sensores automaticamente, através da execução de algoritmos em computadores. "Automático" significa tanto quanto possível substituir o intérprete humano por software e hardware. Isto leva até o conceito de Inteligência Artificial, mesmo que o termo leve possivelmente também a uma falsa esperança, de tal forma que deva ser

melhor descrito como um sistema baseado no conhecimento. Para ambos os casos, a idéia de "conhecimento" é a primordial.

## 2. O Conceito de Conhecimento

Em linguagem coloquial parece claro o que é compreendido como conhecimento. Como exemplo, nos 24 volumes do dicionário alemão do ano de 1901, ("BROCKHAUS") não está inserido o conceito de conhecimento. Quando, entretanto, o conhecimento será processado por um sistema digital, deve ser definido e descrito.

Keith Devlin (1991) define que: "conhecimento engloba um estado de inteligência e uma idéia de verdade"

Esta afirmação descreve por um lado um componente humano do conhecimento, por outro lado um componente absolutamente válido.

Entretanto, este conceito não nos ajuda em muito em nossa discussão sobre inteligência artificial. Makato Nagao (1990) apresenta a fórmula simplificada onde:

$$\text{Conhecimento} = \text{Reconhecimento} + \text{Lógica}$$

Por esta visão está mais claro que o conhecimento deve ser estruturado, que deve ser organizado na forma de *regras*. Isto é uma premissa para que seja possível a comunicação através do conhecimento. Para um sistema de análise de imagens baseado no conhecimento este é um fator essencial.

Trabalhando um pouco mais o conceito de conhecimento, pode-se perceber facilmente que é denotado por duas diferentes naturezas: conhecimento de  *fatos*  e  *regras* . Nem sempre é possível separá-los. Assim o aparato sensorial humano recebe  *fatos*  e os processa diretamente em função dos conhecimento adquiridos, as  *regras* . Ambos os componentes são, sob o ponto de vista do conhecimento, igualmente importantes para a tomada de decisão, contribuindo sinergicamente para a conclusão.

Interessantemente, estes dois lados do conhecimento encontram-se em diversas atividades da vida cotidiana, bem como em distintas especialidades. Isto está apresentado na figura 1 para seis distintas disciplinas. A idéia de fatos e regras são conceitos da inteligência artificial. Em Sistemas de Informações Geográficas (SIG), dados e atributos são fatos e os "modelos de inferência" compõe as regras, das quais serão tiradas as conclusões. Na matemática ou informática são conhecidas linguagens declarativas (PROLOG) e procedurais (FORTRAN). Na jurídica encontram-se por sua vez as evidências e parágrafos das leis. Na lingüística, "conceitos" e "contextos" são, respectivamente, fatos e regras, onde encontram-se os fatos de forma semântica. Até na filosofia, na semiótica (Teoria dos Símbolos), existe representatividade (clássico: Aristóteles, moderno: Frege) e instrumentalidade (clássico: Platão, moderno: Wittgenstein) (Keller, 1995).

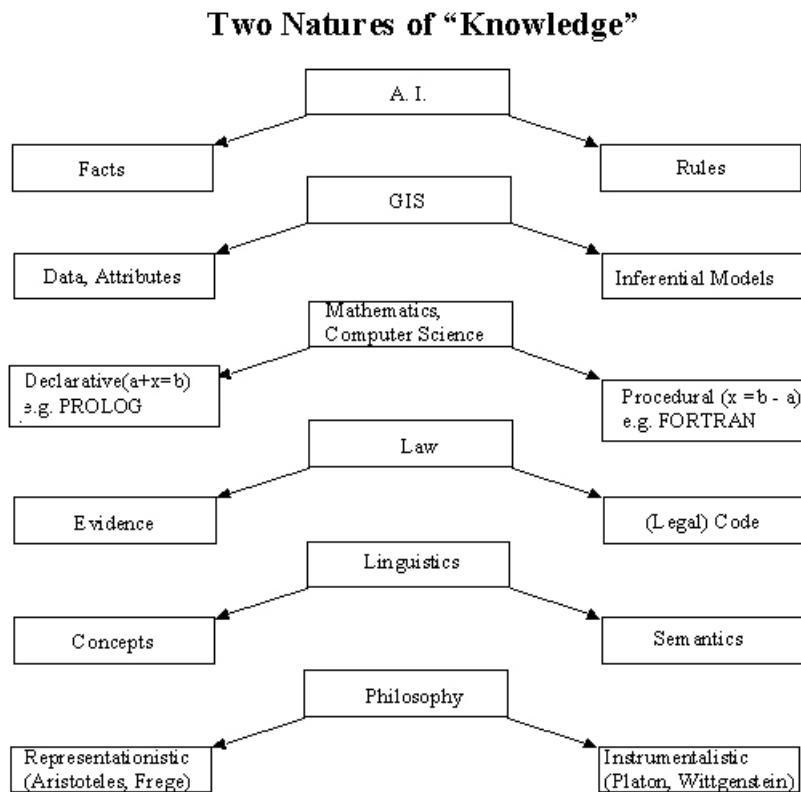


Fig. 1: Fatos e Regras do conhecimento em diferentes especialidades (Bähr, 1998b)

Como exemplo concreto, mostra a figura 2 os dois lados do conhecimento em uma análise via SIG. Trata-se aqui da decisão, se a superfície semiárida do Sahel (África, porção Sul do Saara) é viável para uma forma específica de irrigação. Os dados necessários para tal, os  *fatos* , provêm de fontes de dados como imagens de satélite, trabalhos de campo e cartas topográficas. Estes  *fatos*  são introduzidos em um sistema de  *regras* , o qual pode ser descrito como um modelo, e hierarquicamente conduzidos através de chaves de eliminação. Muitas áreas serão classificadas na classe "não viável" ao passo que outras, viáveis, são classificadas como  *microbacias*  ou como  *macrobacias* .

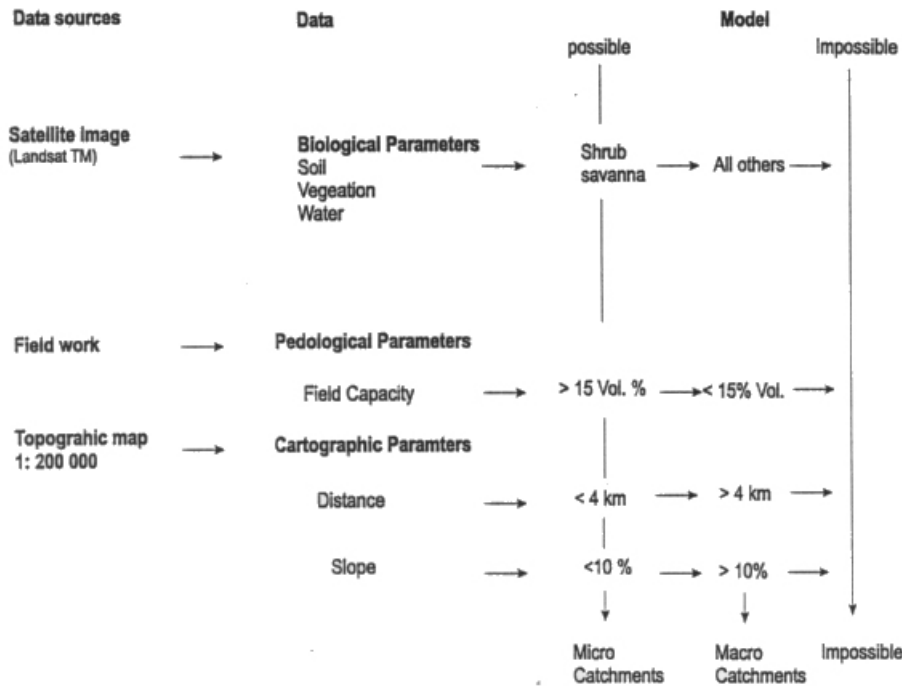


Fig. 2: Componentes do conhecimento incorporados a um SIG-Análise (Tauer e Humborg, 1992; Bähr, 1998)

### 3. Estado Atual da Análise de Dados com Base no Conhecimento, Mostrada Através dos Trabalhos Realizados no IPF/Universidade de Karlsruhe

A figura 3 mostra esquematicamente o fluxograma da análise de imagens com base no conhecimento. A partir dos dados brutos de imagens serão extraídas características, que serão impressos como um gráfico de características (p.ex. "Dados-Spaguetti"). A reunião destas formas primitivas em um objeto pressupõe a descrição deste objeto. Ao final os objetos serão agregados em um nível de processamento superior, o que leva a uma completa descrição da cena. As regras, nas quais esta descrição da cena deve basear-se, devem ser precisamente estabelecidas bem como representar o conhecimento.

O conhecimento pode estar representado no modelo de um nível mais ou menos elevado. Um nível menos elevado trata-se do caso em que ao longo da avaliação dos dados, conhecimentos do operador também são introduzidos (conhecimento implícito) como, p. ex., no caso da classificação multiespectral digital. No segundo caso, também denominado conhecimento explícito, o conhecimento é previamente modelado.

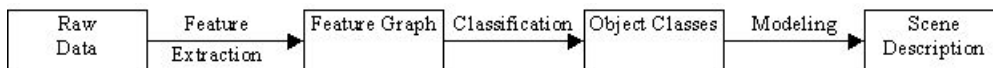


Fig. 3: Fluxograma da análise de imagens

A seguir serão apresentados alguns trabalhos científicos, que foram desenvolvidos no IPF/Universidade de Karlsruhe nos últimos dois anos, avaliando-se em que sentido foram representados e utilizados componentes do conhecimento. A tabela 1 mostra quatro áreas de pesquisa, cuja seqüência vai desde imagens de satélite em pequena escala, passando por fotografias aéreas até cartas vetoriais em grande escala (coluna da esquerda). Os títulos na horizontal seguem o fluxograma da figura 3.

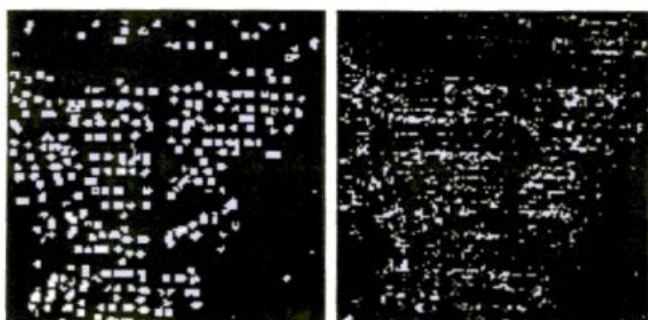
Tabela 1: Áreas de pesquisa, cuja seqüência vai desde imagens de satélite em pequena escala, passando por fotografias aéreas até cartas vetoriais em grande escala.

Project	Input (Raw Data)	Feature Graph	Knowledge Representation	Object Classes Scene Description	Modelling Uncertainty
<b>Object Classification in High Resolution Satellite Imagery</b> (K. Segl) 1996	MOMS-02 Mode 6 (channel 2,3 + 4) + PAN 4,5m resolution	Object features: spectral and multishape	Neuronal Networks:  System training by operator (implicit)	Landuse classes (4) and objects (buildings, streets, trees)	Friedmann-Rafsky-Test for Normal Distribution

<b>Map supported Landuse Classification by Satellite Imagery</b> (K.-J. Schilling et al.) 1996	LANDSAT-TM channels 1,2,3,4,5,7 30m resolution ATKIS DLM 200  (digital map base)	Class features: spectral, textural, boundary, compactness  etc.	Delaunay Triangulation: System training by operator  (implicit)	Landuse classes  Change detection,  Map updating	(under development)
<b>Map supported Object Extraction by Aerial Photography</b> (F. Quint) 1997	German Base Map 1:5000 Aerial Colour Image 1:6300  30 cm resolution	Linear primitives in maps and images  Homogeneous regions in images	Semantic Net  (explicit)	Complete Scene Description  (buildings, parking areas, green, cars)	Extension of Dempster-Shafer Theory Propagation rules for belief
<b>Object Extraction in Large Scale Maps</b> (R. Brügelmann) 1998	German Base Map 1:5000  25cm resol.  (50 &micro;m in Map)	Nodes, lines, regions together with neighbourhood	Markov Random Fields based on Classification  (explicit)	Complete Map Description  (houses, boundaries, symbols, parcels)	All steps accompanied by stochastic models

### 3.1. Classificação de Objetos em Imagens de Satélite de Alta Resolução (Segl, 1996)

O autor utiliza como dados brutos as imagens multiespectrais e pancromáticas do satélite MOMS-02, com uma resolução no solo de 4.5 m. Este trabalho sinaliza a possibilidade de extração não apenas das características espectrais, como também formas. Estas características serão modeladas conjuntamente em uma rede neural. O sistema será treinado pelo operador, tratando-se portanto de conhecimento implícito. A partir da consideração do formato do objeto foi possível a recuperação de pequenos objetos, como casas, grupos de árvores e caminhos, o que não seria possível se fosse considerada unicamente a classificação dos dados multiespectrais.



Esquerda: Segmentação com apoio de rede neural  
Direita: Classificação pelo método da máxima verossemelhança

□ Segmento de Edificação    ■ Indefinido

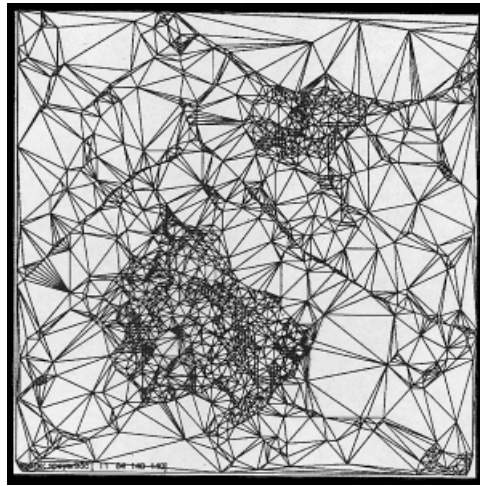
Fig. 4: Segmentação de edifícios a partir de imagens MOMS (Segl, 1996)

A rede neural aplicada por Segl (1996) apresenta a rede neural com camada de entrada, camada interna e camada de saída. Para que fosse possível a combinação de cada sinal de entrada com os demais, Segl agrupou-os através de uma união semântica. A figura 4 mostra o resultado para a classe "casa" no Zimbábue. A comparação entre a classificação com base na máxima verossemelhança e em redes neurais com pré-conhecimento da forma dos objetos mostra claramente a supremacia da técnica desenvolvida. A clara separação entre as "casas" torna possível que também nesta escala já seja possível a contagem do número de edificações e, com isso, indiretamente avaliar-se o número de habitantes. Tal aplicação seria de grande interesse também no Brasil.

### 3.2. Classificação do Uso da Terra a partir de Imagens de Satélite e com Suporte de Cartas (K.-J. Schilling et al., 1996)

Aqui o material de saída está composto de imagens LANDSAT/TM convencionais, com resolução de 30 m. Foi adicionada uma carta vetorial digital existente na escala de 1:200 000. Atualmente as cartas topográficas vetoriais oficiais já encontram-se cada vez mais na forma digital; esta é também a tendência futura. Na Alemanha os dados são armazenados no Sistema Oficial de Informações Topográficas e Cartográficas (ATKIS - "Amtliches Topographisch -und Kartographisches Informations System"). Sugere-se treinar a análise das imagens basicamente com dados da carta. Esta "fusão de dados" descreve um dos pontos fundamentais das atividades de pesquisa no IPF.

A disponibilidade de dados cartográficos na forma digitais facilita a definição de áreas de treinamento para a classificação digital dos dados multiespectrais. Após a classificação serão realizadas análises de vizinhança para cada pixel pertencente a determinada classe. Para isso os pixels de uma determinada classe são reunidos por uma malha triangular segundo as regras da triangulação de Delaunaz. As figuras então resultantes são estatisticamente analisadas em função de sua compactidade, extensão, forma, textura entre outros. Desta forma são consideradas novas informações ("conhecimentos").



**Fig. 5:** Triangulação de Delaunay através do pixel classificado (classe "área urbana") (Schilling e Vögtle, 1996)

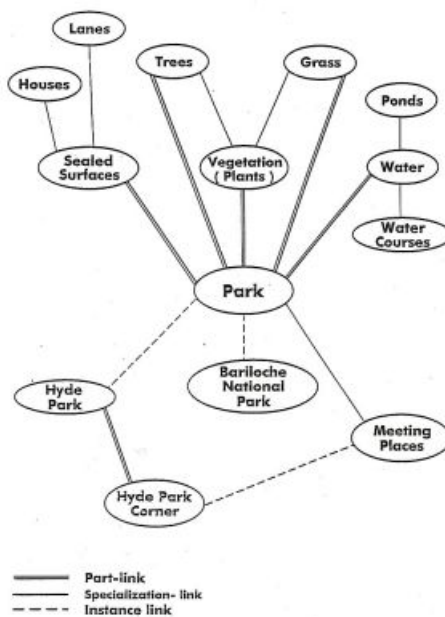
A figura 5 mostra uma parte da imagem LANDSAT/TM após a classificação e triangulação de Delaunay. Depois da análise estatística é relativamente fácil distinguir áreas urbanizadas das típicas formas detidas pelos componentes da rede viária. Os contornos de áreas urbana obtidos na imagem (vilas) são então comparados os contornos destas mesmas áreas em cartas vetoriais oficiais, e a partir da diferença podem ser detectadas mudanças.

Este procedimento é considerado implícito, uma vez que conhecimento adicional foi proporcionado pelo operador ao longo do processo de análise.

### 3.3 Extração de Objetos em Fotos Aéreas com o Apoio de Cartas (F. Quint, 1997)

De forma distinta aos dois exemplos que mostramos anteriormente, esta pesquisa estuda imagens e material cartográfico em grande escala. A resolução no solo para as imagens é de cerca de 30 cm. O gráfico de características compreende primitivas lineares tanto para a carta quanto para a imagem, sendo que para as imagens foram ainda adicionalmente definidas áreas homogêneas. Também aqui trabalha-se novamente de forma a análise conjunta de imagem e carta, procurando aproveitar as vantagens provenientes da união sinérgica destas informações. Uma comparação direta entre imagem e informações cartográficas, assim como ocorre sem problemas para o operador humano, não é possível no caso do processamento computadorizado de imagens. Aqui as bases de dados devem ser transformadas do nível de ícones para o de símbolos. Através destas transformações os dados de ambos as bases são *a priori* formalmente estruturadas e preparadas.

Como modelo para esta estruturação serão utilizadas redes semânticas. A figura 6 mostra um exemplo de uma tal rede semântica. É definida como gráfico e mostra nos nós os conceitos e nas linhas as relações entre os conceitos. O exemplo mostra um caso bastante simples, que em linguagem natural poderia ser traduzido da seguinte forma:



**Fig. 6:** Armazenamento de informações geográficas na forma de uma rede semântica (Bähr e Schwender, 1996)

Vegetação, áreas impermeáveis e água são componentes de parques e através de "links" estão ligados a eles. Árvores e grama também são componentes de parques ("links"). Árvores e grama são casos especiais de vegetação ("links" de especialização). O Parque Nacional de Bariloche e o Hyde Park são instâncias locais de parques ("links" de instância). O Hyde Park é um setor do Hyde Park ("part-link"), é um "ponto de encontro", que detém parte do objeto parque como, p. ex., áreas de lazer. O Hyde Park Corner é uma instância de "Ponto de encontro" determinado e claro ("link" de instância); exemplo de H.-P. Bähr e A. Schwender, 1996.

O armazenamento explícito do conhecimento em uma rede semântica permite uma completa descrição da cena para distintos objetos.

Para cenas mais complexas isto leva a extensas redes de conexões. Por este motivo é indicado o uso de imagens associado a material cartográfico de grande escala, sendo possível a análise de mudanças a partir da determinação das diferenças entre ambos.

### 3.4. Extração de Objetos em Cartas de Grande Escala (R. Brügelmann, 1998)

Na visão do IPF não há nenhuma grande diferença básica entre a análise de imagens de tom médio/multiespectrais e cartas vetoriais. Na prática, a análise de cartas vetoriais é mais fácil, uma vez que as características são inequivocamente passíveis de extração.

R. Brügelmann desenvolveu a tarefa de analisar de forma automática as cartas vetoriais na escala 1:5 000 escanerizadas. Em uma primeira linha tratava-se da segmentação de casas, que nas cartas cadastrais alemão são caracterizadas por hachuras. Para a automação, o primeiro passo necessário foi a digitalização da carta para o formato raster através do escaner. O modelo resultante mostra relações e vizinhanças típicas, as quais são apresentadas no exemplo de uma casa em um limite na figura 7. Como gráfico de características usam-se nós, linhas e áreas conjuntamente com suas vizinhanças. Isto tudo pode ser representado muito bem pelo campo de coincidência de Markoff. O conhecimento também será representado neste caso de forma explícita. O gráfico na figura 7 mostra os elementos e a sua respectiva relação. Como elemento central está o "espaço intermediário" ("ZWI").

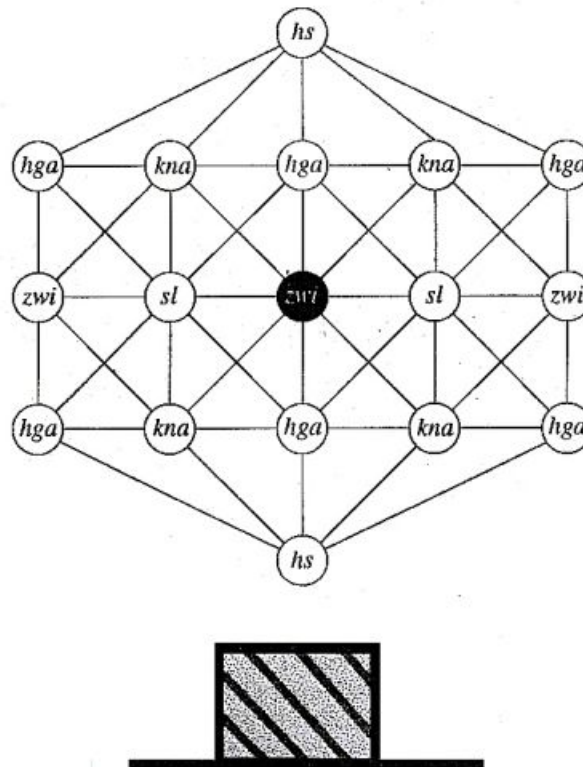


Fig. 7: Estruturação de um modelo de uma planta baixa de uma casa (situada em um limite da carta) através de campos de coincidência de Markoff (R. Brügelmann, 1998)

Existem também vizinhanças de diferentes categorias que, conjuntamente com sua avaliação estocástica, são armazenadas nos campos de coincidência de Markoff. O resultado desta análise é uma descrição completa da carta. A taxa de recuperação de objetos gira em torno de 95%.

## 4. A Importância Substancial do Contexto

Com estas evidências, podemos afirmar que o conhecimento pode ser representado com sucesso por meio de redes. Foram apresentados exemplos de representações implícitas, como redes neurais e triangulação de Delaunay, e explícitas, como redes semânticas e campos de coincidência de Markoff. A base de toda rede é a representação em gráfico, compreendendo nós (detêm o conceito) e linhas (detêm as linhas, relações). Este design encontra-se novamente de acordo com a rede neural humana. Um exemplo pode ser visto na figura 8. Apesar dos itens serem emprestados do alemão, pode-se notar que algumas áreas destas redes neurais são também compreensíveis por pessoas que não falam este idioma (p. ex., Lufthansa/Steward/Anchorage/Alaska/etc). Seria uma análise à parte, trabalhar a questão da mudança de idioma e desta forma não considerar apenas conceitos, mas também as relações entre eles. Somente os conceitos mostram diferenças significativas em distintos idiomas. As relações, no entanto, permanecem imutáveis. A propósito, a rede encontra-se de forma semelhante e um pouco reduzida no conhecido livro de Hofstadter (1978).



**BÄHR, H.-P.:** *GIS-Introduction*. In: BÄHR, H.-P., VÖGTLE, Th. (eds.): *GIS for Environmental Monitoring*. Schweizerbarth Verlag Stuttgart, a ser publicado em 1998.

**BÄHR, H.-P. e SCHWENDER, A.:** *A Linguistic Confusion in Semantic Modelling*. In: KRAUS, K., WALDHÄUSL, P. (eds.): *Proceedings do XVIII Congresso da ISPRS, Viena, Julho/1996*. International Archive of Photogrammetry and Remote Sensing, Vol XXXI, Parte B6, Comissão VI, pp. 13-17.

**BÄHR, H.-P.:** *From Data to Inference: Examples for Knowledge Representation in Image Understanding*" Symposium Comm. III da ISPRS, Columbus/Ohio, 1998b.

**BRÜGELMANN, R.:** *Rasterbasierte Methoden zur Gebäudeextraktion aus gescannte Kanten*. Deutsche Geodätische Kommission, Munique, Série C, a ser publicado em 1998.

**DEVLIN, K. J.:** *Logic and Information*. Cambridge: Cambridge University Press., 1995.

**HOFSTADTER, D. R.:** *Goedel, Escher, Bach: an Eternal Golden Braid*. Nova Yorque: Basic Books, 1979.

**KELLER, R.:** *Zeichentheorie - zu einer Theorie semiotischen Wissens*. Francke Verlag Tübingen und Basel, 1995.

**NAGAO, M.:** *Knowledge and Inference*. Academic Press, 1990.

**QUINT, F.:** *Kartengestützte Interpretation monokularer Luftbilder*. Deutsche Geodätische Kommission da Academia Bavária de Ciência. Dissertação, Série C, caderno 477, Munique, 1997.

**SCHILLING, K.-J. e VÖGTLE, Th.:** *Satellite Image Analysis Using Integrated Knowledge Processing*. In: KRAUS, K., WALDHÄUSL, P. (eds.): *Proceedings do XVIII Congresso da ISPRS, Viena, Julho/1996*. International Archive of Photogrammetry and Remote Sensing, Vol XXXI, Parte B3, Comissão III, pp. 752-757.

**SEGL, K.:** *Integration von Form- und Spektralmerkmalen durch künstliche neuronale Netze bei der Satellitenbildklassifizierung*. Deutsche Geodätische Kommission da Academia Bavária de Ciência. Dissertação, Série C, caderno 468, Munique, 1996.

**TAUER, W. e HUMBORG, G.:** *Runoff Irrigation in the Sahel Zone*. Technical Centre for Agricultural and Rural Cooperation, Margraf Weikersheim, 1992.

**WITTGENSTEIN, L.:** *Philosophical Investigations*. Basil Blackwell Oxford, 1953.

### **Agradecimento**

O autor agradece a Luiz E. Renuncio, doutorando no IPF pelo convênio UFSC/Univ. de Karlsruhe, pela tradução deste texto do alemão para o português.