

Aplicação do programa CORINE e classificação baseada em objetos para mapeamento da cobertura do solo de faixa de domínio

Alessandra Mariane Knopik Beltrame^{1,2}
José Alberto Quintanilha¹

¹ Escola Politécnica da Universidade de São Paulo – POLI/USP
Cidade Universitária, São Paulo - SP, Brasil

² Centro Universitário Senac-SP
Campus Santo Amaro, São Paulo – SP, Brasil
aleknopik@yahoo.com.br
jaquinta@usp.br

Abstract. The principal aim of this paper is the mapping of land cover area of metropolitan railway right of way. The study focuses on the outskirts of the Line F metropolitan train in Sao Paulo. The methodology includes the use of the high spatial resolution satellite image - IKONOS II, object-based classification and application of the CORINE Land Cover Programme. The regions-based classifiers use, besides spectral information of each pixel, the spatial information that involves the relationship between the pixels and its neighbours. Through visual interpretation, were identified twelve classes of land covers, which features guided the segmentation parameters. The segmented image was classified using Bhattacharyya distance and post-classified using rules and criteria established for each class, according to their attributes and contextualization. The performance of the final classification was evaluated based on Kappa coefficient, resulting acceptable level of correspondence. The mapping of land cover generated from the classification, shows that the outskirts of the metropolitan railway's right way is predominantly composed by impermeable materials - such as asphalt and concrete - but has significant presence of natural elements - such as vegetation and water.

Palavras-chave: remote sensing, object-based classification, railway right of way, CORINE, sensoriamento remoto, classificação baseada em objetos, faixa de domínio de ferrovia.

1. Introdução

O volume de investimentos realizado pelas concessionárias de ferrovias no setor cresceram 215% entre 2003 e 2005, chegando a marca de R\$3,37 bilhões no ano de 2005 (DNIT, 2008). Apesar dos investimentos, a recuperação da malha ferroviária atual encontra-se entre os principais desafios. Em dezessete estados, 434 leitos ferroviários foram tomados por ocupações irregulares, como favelas, hotéis, clubes, marinas e loteamentos de luxo. Ocupações irregulares ocorrem também ao longo do eixo de linhas férreas ativas, caracterizadas por crescentes invasões das faixas de domínio.

A gestão da faixa de domínio é significativamente onerosa devido à considerável extensão da malha ferroviária brasileira – 29 mil quilômetros - aliada à diversidade geográfica. Neste aspecto, o sensoriamento remoto apresenta-se como uma ferramenta eficiente no monitoramento de corredores de transporte, e suas respectivas áreas circunvizinhas.

O objetivo deste trabalho é o mapeamento da cobertura do solo ao longo da faixa de domínio de trem metropolitano e suas adjacências, baseado no sistema CORINE - Coordination of Information of Environment (Bossard et al., 2000) e técnica de classificação baseada em objetos.

O CORINE é um sistema de classificação de cobertura do solo, tendo sido criado com objetivo de fornecer informações sobre o meio ambiente à União Européia. A caracterização das classes é feita com base em suas composições e arranjos espaciais. A classificação segue uma ordem hierárquica, sendo o primeiro nível composto por 5 classes, o segundo por 15 e o terceiro por 44 classes. A partir do quarto nível o sistema encontra-se em aberto, cabendo ao

analista definir as demais classes, de acordo com os materiais identificados na cena trabalhada.

A classificação baseada em objetos apresenta-se como um método alternativo aos métodos tradicionais de classificação, pois utiliza tanto a informação espectral de cada pixel como também a informação espacial que envolve a relação entre os pixels e sua vizinhança. Estes classificadores procuram simular o comportamento de um foto-intérprete, ao reconhecer áreas homogêneas de imagens, baseados nas propriedades espectrais e espaciais de imagens. (INPE, 2007).

Um dos algoritmos utilizados na classificação baseada em objeto é o Bhattacharyya, que estima a separabilidade estatística para duas classes espectrais, ou, mais corretamente, mede a probabilidade de uma classificação mais correta (Mather, 2005). Para cada objeto, ou região, é calculada a distância de Bhattacharyya entre duas classes, ω_1 e ω_2 , conforme equação 1. A distância é computada para cada par de classe dado um número m de feições. Todas as possíveis combinações são testadas e, ao término do processo, a região avaliada será atribuída à classe que apresentar menor distância em relação à outra classe (Mather, 2005).

A classificação baseada em objeto é antecedida pela segmentação, que consiste na divisão de uma imagem em regiões que melhor representam os objetos de interesse numa cena, sendo que os atributos de cada região, como área, formato, parâmetros estatísticos e texturas podem ser extraídos e adicionados na análise dos dados (Bins et al., 1996). Como os objetos identificados numa imagem podem ser delineados pela segmentação, o número de elementos a serem classificados posteriormente é reduzido significativamente. Desta forma, a qualidade da classificação está relacionada diretamente à qualidade da segmentação (Meinel e Neubert, 2002).

A disponibilidade cada vez maior de dados de SIG abre novas e potenciais opções para avaliação de dados de sensoriamento remoto (Schiewe e Tufle, 2006). Conforme Blaschke et al. (2007) as funcionalidades do SIG podem ser integradas diretamente no processo de classificação da imagem, fornecendo informações relevantes sobre os objetos ou regiões identificados na imagem, de forma a garantir que a classificação final tenha um resultado melhor do que se utilizasse apenas informações espectrais.

2. Materiais

2.1 Área de estudo

A área de estudo concentra-se sobre um trecho de 8,5 km de extensão da Linha F da CPTM (Companhia Paulista de Trem Metropolitano), limitada pelas estações Comendador Ermelino e Itaim Paulista, localizada entre as latitudes 23°28'19"S e 23°30'35"S e longitudes 46°29'04"W e 46°23'43"W, em São Paulo, capital. A região é densamente povoada, inclusive com presença de favelas juntas à linha férrea, mas também apresenta elementos bem diversificados, como rios, vegetação densa e galpões industriais.

2.2 Materiais

Utilizou-se imagens IKONOS II ortorretificada, datada de outubro de 2002, com fusão da banda pancromática e de combinações coloridas das bandas NIR, Red, Green e Blue (infravermelho próximo, vermelho, verde e azul, respectivamente), para a composição da cena multispectral com resolução espacial de 1m.

Bases de dados digitais de Quadras e Logradouros do município de São Paulo, integrantes do programa São Paulo Protege (http://portal.prefeitura.sp.gov.br/secretarias/assistencia_social/menu_secretaria/sp_protege), fornecidas ao Laboratório de Geoprocessamento. As duas bases estão em escala 1:2.000 e sistema de projeção Latitude / Longitude.

Para fins de processamento e análise, utilizou-se os programas ENVI 4.2, Spring 4.2 e ArcGIS 9.1.

3. Metodologia

A metodologia proposta baseou-se em quatro processos macros, descritos a seguir.

3.1 Pré-processamento

Para obtenção de uma única imagem que cobrisse a Linha F em toda sua extensão foram necessárias 4 cenas IKONOS II originais, que foram mosaicadas e balanceadas. A imagem resultante apresentou tamanho consideravelmente grande, inviabilizando o seu processamento digital.

Para fins de processamento, delimitou-se uma área de 750m ao redor da linha, passando a ser o objeto de estudo. Essa distância foi definida considerando-se a melhor combinação possível entre uma área capaz de incluir todos os elementos visualmente identificados na cena e o tamanho final da imagem. A área final de estudo pode ser visualizada na Figura 1.

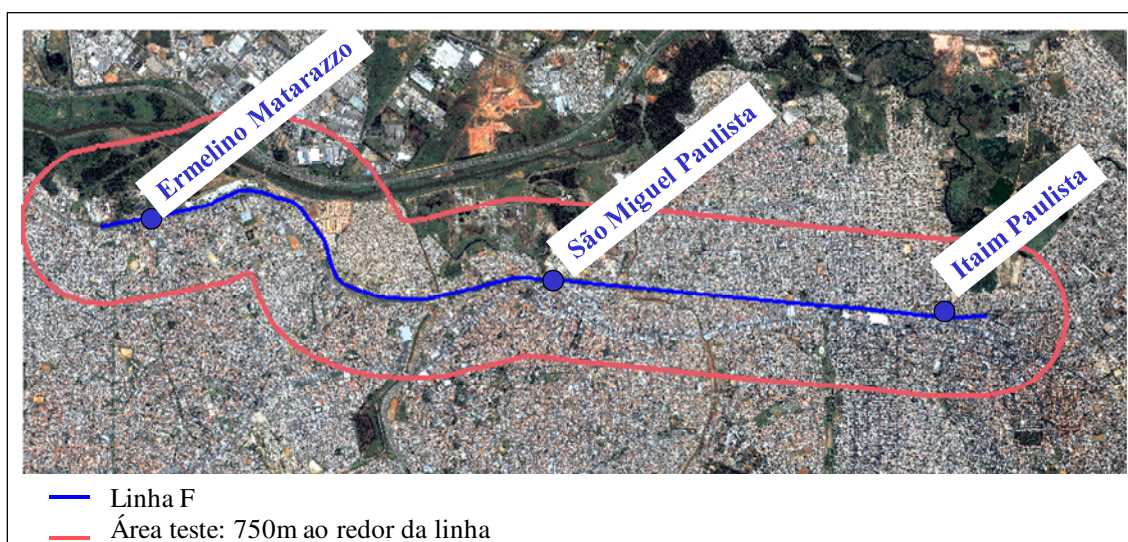


Figura 1. Delimitação da área de estudo, compreendendo 750m ao redor da linha de trem.

3.3 Processamento

3.3.1 Definição das classes de objetos

A definição das classes de objetos foi feita com base no sistema CORINE, a partir do nível III, classe 1.2.2, correspondendo a estradas, ferrovias e faixas de servidão. A interpretação visual da imagem permitiu identificar objetos e materiais com maior representatividade na cena, além de suas características, como cor, forma e textura. As classes identificadas foram relacionadas no nível IV, conforme Quadro 1.

Quadro 1. Sistema de classificação CORINE para cobertura do solo.

Nível I	Nível II	Nível III	Nível IV
1. Áreas artificiais	1.1 Tecido urbano	1.1.1 Tecido urbano contínuo	
		1.1.2 Tecido urbano descontínuo	
	1.2 Comércio, indústria e transporte	1.2.1 Unidades industriais e comerciais	
		1.2.2 Estradas, ferrovias e faixas de servidão	1.2.2.1 Vegetação densa
			1.2.2.2 Vegetação rasteira

			1.2.2.3 Água
			1.2.2.4 Solo exposto claro
			1.2.2.5 Solo exposto médio
			1.2.2.6 Solo exposto escuro
			1.2.2.7 Cobertura metálica
			1.2.2.8 Amianto claro
			1.2.2.9 Amianto escuro
			1.2.2.10 Asfalto
			1.2.2.11 Telha cerâmica
			1.2.2.12 Sombra
			1.2.3 Áreas portuárias
			1.2.4 Aeroportos
1.3 Minas, aterros e canteiros de obra	1.3.1 Extração mineral		
	1.3.2 Aterros		
	1.3.3 Canteiro de obras		
1.4 Áreas verdes artificiais	1.4.1 Vegetação urbana		
	1.4.2 Áreas de esporte e lazer		

Fonte: Adaptado de Bossard *et al.* (2000).

3.3.2 Segmentação

A segmentação foi realizada para dividir a imagem de acordo com as feições dos materiais identificados na interpretação visual. Para tanto, usou-se a segmentação por crescimento de regiões, agrupando os pixels em regiões, anexando a cada pixel inicial outros pixels que apresentaram propriedades similares, como nível de cinza, textura e cor. Os parâmetros utilizados foram 30 para similaridade e 50 para área.

3.3.3 Classificação da imagem

A imagem foi classificada em regiões, que por sua vez, correspondem aos polígonos gerados na etapa da segmentação. A classificação foi realizada usando-se o algoritmo Bhattacharyya, disponível no Spring 4.2, tendo como contexto as quatro bandas da imagem.

3.3.4 Pós-classificação

Nessa fase fez-se a identificação e aplicação de critérios, ou regras, sobre a imagem já classificada, de forma a corrigir possíveis erros cometidos na etapa anterior. Os polígonos classificados corretamente foram analisados tanto com base em seus atributos, como área e perímetro, como também em seu contexto na cena, servindo de base para a caracterização das classes. Critérios de reclassificação também foram gerados a partir das relações topológicas dos polígonos da imagem classificada, com os das bases digitais de quadras e logradouros e também com geração de imagem NDVI.

3.3.4.1 Asfalto

Para melhor diferenciar esta classe da de amianto escuro, definiu-se como asfalto todo polígono de área igual ou menor a 2.200m^2 - medida correspondente à média da área dos polígonos de referência da classe asfalto - e que interceptasse o eixo de logradouros.

3.3.4.2 Telha cerâmica x solo exposto médio

Para reduzir a confusão entre as duas classes, definiu-se solo exposto médio todo polígono classificado como telha cerâmica que estivesse fora dos limites das quadras. Uma segunda regra foi aplicada na seqüência, segundo a qual polígonos classificados como solo

exposto médio, cuja razão entre área e perímetro estivesse no intervalo de 0,35 a 0,49 deveria ser classificado como telha cerâmica.

3.3.4.3 Água

Definiu-se como classe água todo polígono classificado como asfalto e amianto escuro em adjacência com polígonos da classe vegetação rasteira e cujo perímetro apresentasse valores entre 500m e 600m.

3.3.4.4 Vegetação densa

Para melhorar o resultado desta classe utilizou-se uma imagem NDVI, tendo sido aplicada a seguinte regra: polígonos atribuídos às classes asfalto e amianto escuro contidos em áreas com valor de NDVI acima de 18, numa escala de 0 a 100, pertencem à classe vegetação densa.

3.4 Avaliação da classificação final

Os polígonos de referência tiveram a veracidade de suas respectivas classes obtida através da interpretação visual da imagem IKONOS II utilizada no projeto. A ausência de uma visita a campo se justificou, principalmente, pelo considerável intervalo de tempo decorrido entre a geração da imagem IKONOS II (2002) com um possível levantamento de campo atual.

Tradicionalmente a área total das classes é obtida pela somatória dos pixels correspondentes na imagem. Neste trabalho, contudo, a contagem de pixels foi substituída pelo somatório das áreas dos polígonos, sendo o m^2 sua unidade de medida.

Para avaliação da classificação gerou-se a matriz de confusão ou matriz de erro, que fornece a distribuição das unidades classificadas correta e erroneamente, e a aplicação do índice Kappa, que consiste na medida da diferença de concordância dos dados de referência e a classificação automática e a probabilidade de concordância entre os dados de referência e a classificação aleatória (Lillesand *et al.*; 2004).

4. Resultados

4.1 Segmentação

A combinação dos parâmetros 30-50, similaridade e área respectivamente, resultou numa segmentação consideravelmente detalhada, eficiente na delimitação dos objetos de menores dimensões da cena, mas não tão eficiente em se tratando de elementos de maior dimensão, como parques e rios, por representá-los em número de segmentos muito superior ao que seria necessário. Considerando-se que a forma dos objetos não foi utilizada como critério de classificação, optou-se por manter esses valores.

4.2 Classificação por regiões

Comparando-se visualmente a imagem original com a classificação, percebe-se uma confusão entre algumas classes, como solo exposto médio e telhado cerâmico, asfalto com concreto médio e escuro e água e sombra. Esta confusão já era esperada devido à semelhança do comportamento espectral verificada entre os elementos constituintes destas classes (Souza e Kux, 2005).

Para melhor estimar o desempenho desta classificação, bem como quantificar o grau de confusão existente e entre quais classes, foi gerada uma matriz de confusão permitindo essas e outras análises, representada pela Tabela 1.

Tabela 1. Matriz de confusão da classificação por regiões.

	Água	Amianto claro	Amianto escuro	Asfalto	Cobertura metálica	Solo exp claro	Solo exp médio	Solo exp escuro	Sombra	Telha cerâmica	Veg densa	Veg rasteira
Água	32%	-	60%	-	-	-	-	-	1%	-	6%	-
Amianto claro	-	92%	2%	-	1%	4%	-	-	-	1%	-	-
Amianto escuro	-	1%	92%	5%	-	-	-	1%	-	-	-	-
Asfalto	-	-	28%	72%	-	-	-	-	-	-	-	-
Cobertura metálica	-	39%	-	-	61%	-	-	-	-	-	-	-
Solo exp claro	-	4%	-	-	-	88%	-	1%	-	6%	-	-
Solo exp médio	-	-	-	-	-	8%	42%	23%	-	27%	-	-
Solo exp escuro	7%	-	3%	-	-	-	7%	82%	-	2%	-	-
Sombra	6%	-	2%	6%	-	-	25%	2%	52%	2%	5%	-
Telha cerâmica	-	1%	-	1%	-	1%	31%	2%	-	65%	-	-
Veg densa	11%	-	11%	-	-	-	-	7%	3%	4%	44%	20%
Veg rasteira	2%	-	-	-	-	-	-	17%	0%	2%	3%	77%
Exatidão usuário	37%	84%	88%	79%	94%	84%	49%	30%	67%	62%	65%	96%
Exatidão geral	79,25%											

A exatidão geral foi de 79,25%, considerada boa. Todavia, não reflete uma situação de bom desempenho em todas as classes, algumas classes muito bem avaliadas compensaram o baixo desempenho de outras. Como a finalidade do presente trabalho é o mapeamento da cobertura do solo, todas as classes possuem a mesma importância e devem apresentar um resultado final de classificação razoável, o que foi obtido na etapa seguinte.

4.3 Critérios de reclassificação

4.3.1 Asfalto x amianto escuro

A classe asfalto apresentou ganho de 7%, mas amianto escuro caiu 10%, mas mesmo com exatidão de 82% permanece bem avaliado. A exatidão do usuário teve uma queda considerável em virtude de asfalto assumir para si parte do erro de comissão anteriormente da atribuída à classe amianto escuro. A exatidão geral apresentou ligeira queda, ocasionada pela queda na avaliação do amianto escuro. Todavia, como esperava-se melhorar o índice de exatidão geral com a correção de outras classes, e conseguiu-se melhorar o desempenho da classe asfalto, o critério foi considerado válido.

4.3.2 Telha cerâmica x solo exposto médio

A aplicação da primeira regra promoveu uma melhora de 42% para 62% de desempenho da classe solo exposto médio. Com aplicação do segundo critério a avaliação da classe telha cerâmica melhora consideravelmente, passando de 65% para 82%, mas a classe solo exposto médio volta ao patamar anterior de 46%. A exatidão geral permaneceu quase inalterada. Devido aos novos resultados obtidos para telha cerâmica, optou-se por manter esse critério, mesmo com a desvalorização do solo exposto médio. Considerou-se também que, embora o presente trabalho não pretenda priorizar nenhuma classe, solo exposto é uma categoria de cobertura do solo mais suscetível à mudanças do que telha cerâmica. Desta forma, acredita-se que um mapeamento com maior precisão para classe telha seria de maior valia do que se mantivesse os índices de 62% para a classe solo exposto médio.

4.3.3 Água

Obteve-se uma melhora significativa no desempenho da classe água, tanto na exatidão do produtor, de 32% para 76%, como na exatidão do usuário, de 37% para 58%. Uma sensível melhora pôde ser observada também na exatidão do usuário das classes amianto escuro e asfalto. A exatidão geral teve um aumento de quase 1%.

4.3.4 Vegetação densa

A aplicação deste critério eliminou o erro de comissão das classes asfalto e amianto escuro em relação a vegetação densa, resultando no ganho da exatidão do usuário destas classes para 58% e 93%, respectivamente. Já a classe vegetação densa teve uma significativa melhora, passando de 44% para 55%.

4.4 Resultado final da classificação

Uma amostra da classificação final pode ser visualizada na Figura 2.



Figura 2. Um trecho da área classificada após aplicação dos critérios de reclassificação.

Após a aplicação de todos os critérios de reclassificação, chegou-se a um resultado final de 0,74 para o coeficiente Kappa. Os demais índices podem ser verificados na Tabela 2.

Tabela 2. Matriz de erro da classificação final.

	Água	Amianto claro	Amianto escuro	Asfalto	Cobertura metálica	Solo exp claro	Solo exp médio	Solo exp escuro	Sombra	Telha cerâmica	Veg densa	Veg rasteira
Água	76%	-	-	16%	-	-	-	-	1%	-	6%	-
Amianto claro	-	92%	1%	1%	1%	4%	-	-	-	1%	-	-
Amianto escuro	-	1%	84%	15%	-	-	-	1%	-	-	-	-
Asfalto	-	-	21%	79%	-	-	-	-	-	-	-	-
Cobertura metálica	-	39%	-	-	61%	-	-	-	-	-	-	-
Solo exp claro	-	4%	-	1%	-	87%	2%	1%	-	5%	-	-
Solo exp médio	-	-	-	-	-	8%	46%	23%	-	23%	-	-
Solo exp escuro	7%	-	2%	1%	-	-	2%	82%	-	6%	-	-
Sombra	6%	-	2%	6%	-	-	25%	2%	52%	2%	5%	-
Telha cerâmica	-	1%	-	1%	-	1%	14%	2%	-	82%	-	-
Veg densa	11%	-	-	-	-	-	-	7%	3%	4%	55%	20%
Veg rasteira	2%	-	-	-	-	-	-	17%	-	2%	3%	77%

5. Conclusões

A classificação final apresentou um índice de desempenho satisfatório, permitindo identificar e mensurar a presença dos diversos materiais encontrados às margens da linha ferroviária. A classificação inicial, produto do algoritmo Bhattacharyya, apresentou um resultado acima das expectativas, podendo ser melhorado com aplicação de regras e critérios para identificação das classes.

As classes identificadas na cena permitiram a criação do sub-nível IV para o sistema CORINE. Quanto à caracterização das classes, estas podem ser agrupadas em três categorias distintas, de acordo com seu grau de identificação: a) as facilmente caracterizadas apenas pela sua resposta espectral, como amianto claro, vegetação rasteira e solo exposto escuro; b) as que demandaram análise mais apurada de seus atributos e sua contextualização na cena para refinamento de sua identificação, como asfalto, amianto escuro, telha cerâmica e solo exposto médio e, c) as que não se conseguiu delimitar com base apenas nos atributos disponíveis, como cobertura metálica e sombra.

A diversidade de arranjos urbanos encontrados na área do projeto – parque ecológico, indústrias, residências e ocupações irregulares – foi um dos pontos atrativos para o desenvolvimento do trabalho. Entretanto, num trabalho de classificação da cobertura do solo, essa variedade dificultou a criação de regras ou critérios aplicáveis satisfatoriamente à toda região.

O mapeamento da cobertura do solo numa área limite em torno da linha férrea pode ser de grande valia para as concessionárias de ferrovias. Embora o presente trabalho não tenha se prestado a essa finalidade, às diversas classes de cobertura podem ser atribuídos pesos de forma a classificar as regiões também com relação ao seu grau de risco para a ferrovia.

Bibliografia

BOSSARD, M.; FERANEC, J.; OTAHEL, J. CORINE land cover technical guide – Addendum 2000. **Technical Report**, nº 40, European Environment Agency, 2000.

BINS, L.S.; FONSECA, L. M. G.; ERTHAL, G. J.; MITSUO, F. Satellite imagery segmentation: a region growing approach. In: VIII Simpósio Brasileiro de Sensoriamento Remoto, 1996, Salvador. **Anais...** São José dos Campos: INPE, 1996. p. 677 - 680.

BLASCHKE, T.; GLÄSSER, C.; LANG, S. Processamento de imagens num ambiente integrado SIG/sensoriamento remoto: Tendências e conseqüências. In: **Sensoriamento Remoto e SIG avançados**. Ed. BLASCHKE, T.; KUX, H. 2 ed. São Paulo: Oficina de Textos, 2007. p. 11-18.

DNIT – DEPARTAMENTO NACIONAL DE INRA-ESTRUTURA DE TRANSPORTE. Disponível em <http://www.dnit.gov.br/noticias/investimentoferroviario>. Acessado em 07/01/2008.

INPE – INSTITUTO NACIONAL DE PESQUISAS ESPECIAIS. SPRING 4.2: sistema de ajuda on-line do Spring – manual do usuário. Disponível em <http://www.dpi.inpe.br/spring/portugues/tutorial>. Acesso em 17/01/2007.

MATHER, P. M. **Computer Processing of Remotely-Sensed Images**. 3º Ed, Chichester: John Wiley & Sons Ltd, 2005. 324p.

MEINEL, G.; NEUBERT, M. A comparison of segmentation programs for high-resolution remote sensing data. In: **XXth ISPRS – International Society for Photogrammetry and Remote Sensing Congress**, Istanbul, 2004.

SCHIEWE, J.; TUFTE, L. O potencial de procedimentos baseados em regiões para a avaliação integrada de dados de SIG e sensoriamento remoto. In: **Sensoriamento Remoto e SIG avançados**. Ed. BLASCHKE, T.; KUX, H. 2 ed. São Paulo: Oficina de Textos, 2007. p. 56-65.