

INCORPORAÇÃO DE DADOS AUXILIARES PARA MELHORAR A CLASSIFICAÇÃO DA VEGETAÇÃO

¹Aguinaldo Araújo Silva Filho

²Jorge Silva Centeno

³Ana Lúcia Bezerra Candeias

- 1 Sec. de Ciência, Tec. e Meio Ambiente - aguinaldoaraujo@zipmail.com.br
- 2 Universidade Federal do Paraná - centeno@geoc.ufpr.br
- 3 Universidade Federal de Pernambuco - analucia@ufpe.br

RESUMO

O trabalho tem como objetivo estudar a integração de dados auxiliares para melhorar a classificação digital da imagem do satélite Landsat na diferenciação de tipos florestais. Sendo os dados auxiliares relacionado com a topografia do lugar. A área de estudo escolhida localiza-se entre os paralelos de coordenadas 08° 00' 00" e 08° 29' 44" Sul e os meridianos de coordenadas 36° 29' 13" e 36° 29' 26" Oeste de Wgr. Abrange parte dos municípios de Belo Jardim, Brejo da Madre de Deus e Tacaimbó, estes situados no Agreste pernambucano. A metodologia consiste em classificar a imagem que recobre a área de estudo pelo método supervisionado utilizando o algoritmo da máxima verossimilhança ; integrar dados auxiliares antes, durante e depois do processo de classificação da imagem; e analisar os resultados que forem gerados. Para isto, os resultados da aplicação das técnicas de estratificação geográfica, incorporação dos dados como banda auxiliar durante a classificação e a reclassificação como pós processamento são comparados e discutidos. O algoritmo de classificação escolhido é o da máxima verossimilhança gaussiana, disponível no sistema SPRING de processamento de imagens. Como primeiro experimento, a imagem resultante da composição dos canais 5, 4, 3 foi classificada pelo processo supervisionado utilizando-se o algoritmo da Máxima Verossimilhança (Maxver). O mapa temático resultante foi avaliado através de amostras de verificação, sendo verificado um índice de confusão média de 10.35%. Num segundo experimento, efetuou-se a classificação utilizando todas as variáveis, isto é, a imagem multiespectral e os dados de altitude, representados como uma quarta banda. O resultado obtido revela que a confusão média entre classes diminuiu significativamente de (1.11%), principalmente no topo dos morros, onde a vegetação do tipo arbóreo é mais frequente.

Palavras-chave: Classificação; Topografia; Dados auxiliares.

ABSTRACT

The work studies the integration of auxiliary data to improve the digital classification of the image of the satellite Landsat in the differentiation of forest types. Being the auxiliary data related with the topography of the place. The study area chosen he/she is located among the parallel of coordinates $08^{\circ} 00' 00''$ and $08^{\circ} 29' 44''$ South and the meridians of coordinates $36^{\circ} 29' 13''$ and $36^{\circ} 29' 26''$ West of Wgr. This area is a part of the municipal districts of Belo Jardim, Brejo da Madre de Deus and Tacaimbó, these placed in the Rural from Pernambuco. The methodology consists of classifying the image that recovers the study area for the supervised method using the algorithm of the maximum likelihood; to integrate auxiliary data before, during and after the process of classification of the image; and to analyze the results that be generated. For this, the results of the application of the techniques of geographical bedding, incorporation of the data as auxiliary band during the classification and the reclassification as powders processing is compared and discussed. The algorithm of chosen classification is it of the gaussian maximum likelihood, available of the system SPRING of processing of images. As first experiment, the resulting image of the composition of the channels 5, 4, 3 were classified by the supervised process being used maximum likelihood algorithm. The resulting thematic map was evaluated through verification samples, being verified an index of medium confusion of 10.35%. In a second experiment, he occurred the classification using all the variables, that is, the image multiespectral and the altitude data, represented as a fourth band. The obtained result reveals that the medium confusion among classes decreased significantly of (1.11%), mainly in the top of the hills, where the vegetation of the arboreal type is more frequent.

Key-words: Classification; Topography; Ancilares data.

1. INTRODUÇÃO

A vegetação de regiões com clima semi-árido é caracterizada por ter a fenologia de grande parte de sua flora controlada pela disponibilidade de água, que por sua vez apresenta grande variabilidade espacial e temporal. Estas características levam freqüentemente a situações de redução da capacidade de suporte dos ecossistemas semi-áridos que se não forem acompanhados de controles na utilização destes ecossistemas, podem desencadear processos de degradação muitas vezes de recuperação bastante lenta. É esta vegetação que nos períodos de estiagem

representa uma alternativa de fonte de renda para os produtores rurais. Portanto, necessita-se desenvolver técnicas que represente a realidade mais próxima do terreno para que se possa monitorar o desenvolvimento vegetativo das coberturas vegetais e as áreas de cobertura vegetal das regiões semi-áridas para melhor administrá-las.

A análise digital de dados, mais especificamente, imagens digitais de sensoriamento remoto orbital, possibilitou, nos últimos vinte e cinco anos, um grande desenvolvimento das técnicas voltadas para a análise de dados multidimensionais, adquiridos por diversos tipos de sensores. Estas técnicas têm recebido o nome de processamento digital de imagens.

A classificação automática e semi-automática de imagens de satélites vem facilitando diversas áreas do conhecimento. Em nosso caso desejamos classificar melhor dados de vegetação a partir de uma imagem de satélite. Como a imagem traz a informação bidimensional do lugar nem sempre tem-se uma real classificação, então é necessário desenvolver técnicas para incorporar dados auxiliares no processo de classificação da imagem de forma que permita um mapa mais próximo da realidade do terreno.

O maior problema da classificação, de acordo com (Crosta, 1992), é que ela representa uma simplificação bastante grande em relação à enorme complexidade existente na cena visto que ela possui uma dependência da resposta espacial e espectral. O que ocorre é que as classes pré-definidas para a classificação são normalmente uma simplificação em relação à realidade terrestre.

A mistura extremamente variável de uma série de objetos reais, incluídos em um mesmo pixel, irão influenciar também no resultado da classificação. Sendo assim, pode ser cometido erros ao se definir classes que não existam, pois o nível de cinza resultante pode não representar aquele dado tema.

Portanto faz-se necessário a verificação dos resultados das classificações em relação a dados conhecidos do terreno. Então, quaisquer tipos de informações sobre os objetos e superfícies a serem classificados serão auxílios importantes à classificação.

Outro problema da classificação é em relação à maioria das técnicas de processamento da imagem disponíveis, que são baseadas apenas no agrupamento de valores de intensidade espectral, representados pelos níveis de cinza presentes na imagem. Ocorre que, as técnicas padrões só dão conta das características espectrais, desprezando as características texturais e outras.

Uma forma de melhorar a classificação é incluir algumas variáveis auxiliares para estratificar uma determinada classe. Um exemplo seria usar a altimetria para estratificar a vegetação na área em estudo. Já que ela é uma variável importante no porte, tipo e densidade da vegetação.

Outras abordagens sobre a melhoria da classificação seria incluir dados não espectrais auxiliares quando se faz classificação supervisionada, não supervisionada

e/ou classificação fuzzy para extrair a informação desejada. Existe uma variedade de métodos, incluindo o uso de estratificação geográfica, lógica de classificação em camadas e sistemas especialistas (Strahler *et al.*, (1978) ; Hutchinson, (1982)).

A partir do exposto acima, vários cientistas (por exemplo: Strahler *et al.*, (1978); Hutchinson, (1982)), melhoram a acurácia e a qualidade da classificação derivada do sensoriamento remoto pela incorporação de dados auxiliares no processo de classificação.

Várias abordagens podem ser utilizadas para incorporar dados auxiliares no processo de classificação de imagens, de modo a melhorar os resultados. Entre elas a incorporação de dados antes, durante ou após a classificação através de estratificação geográfica, operações de classificadores e/ou escolha pós classificação (Hutchinson, 1982).

Diante do contexto, nesse trabalho utiliza-se a altimetria como dado auxiliar por haver disponibilidade desse dado em forma digital e também por esse dado estar diretamente ligado a ocorrência de tipos florestais nesta região. Nada impede que se use outra variável, por exemplo o solo como dado auxiliar, uma vez que este exerce uma forte influencia na ocorrência de tipos florestais.

Assim, o presente trabalho descreve o resultado de experimentos a respeito do uso da altimetria como variável auxiliar no processo de classificação de imagens Landsat. Para isto, um modelo digital do terreno foi gerado com a mesma resolução das imagens multiespectrais.

2. MATERIAIS E MÉTODOS

2.1 ÁREA DE ESTUDO

A área de estudo foi escolhida por se ter um grande conhecimento da mesma, onde o autor realizou mapeamento da vegetação, estudos batimétricos no açude Bitury, levantamento sócio-econômico, participou no inventário florestal e levantamento de consumo de produtos florestais. Esta localiza-se entre os paralelos de coordenadas 08° 00' 00" e 08° 29' 44" Sul e os meridianos de coordenadas 36° 29' 13" e 36° 29' 26" Oeste de Wgr. Abrange parte dos municípios de Belo Jardim, Brejo da Madre de Deus e Tacaimbó, estes situados no Agreste pernambucano, Figura 01. Está área fica situada na zona de transição entre duas regiões com situações geográficas distintas, o Litoral / Mata e o Sertão, refletida pela umidade e o relevo. Apresenta um tipo climático característico (Tropical Sub-úmido), de temperaturas médias anuais entre 22° e 24°C. As condições climáticas são muito heterogêneas, uma vez que o relevo é muito movimentado e tem forte influência, não só sobre a temperatura, como também sobre a umidade relativa do ar e o regime pluviométrico. O relevo se constitui de uma sucessão de degraus e superfícies aplainadas, com variações altimétricas que chegam a atingir altitudes superiores a

800 metros. Contudo, nas áreas de maiores altitudes e expostas aos ventos alísos de sudeste, aparecem os “Brejos de Altitude”, originalmente cobertos por matas serranas e nas altitudes inferiores a 800 metros aparece uma vegetação arbustiva arbórea aberta ou fechada. A vegetação predominante é a Caatinga hipo e hiperxerófila, cujas espécies apresentam perda das folhas, árvores retorcidas, dotadas de espinhos e a presença de Cactáceas e Bromeliáceas.

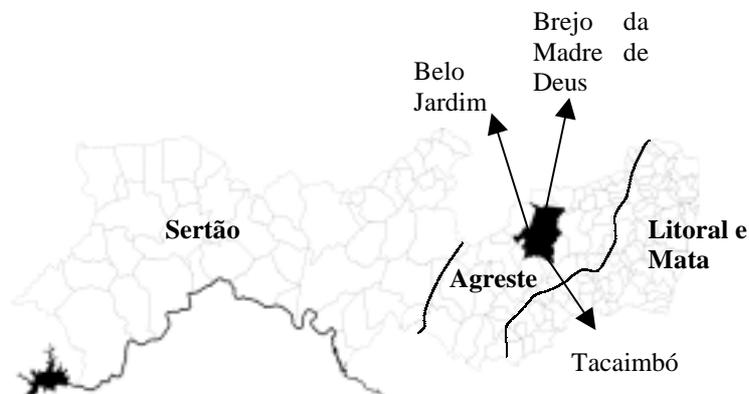


FIGURA 01- MAPA DE LOCALIZAÇÃO DA ÁREA DE ESTUDO

Conforme Silva Filho (1993), o projeto PNUD/FAO/IBAMA/Governo de Pernambuco mapeou a vegetação lenhosa desta região em 1993, por meio da interpretação visual de imagem de satélite (Landsat – 5, TM), escala 1 : 100.000, composição colorida dos canais 5, 4 e 3 em papel fotográfico. Este mapeamento permitiu identificar 3 tipos florestais em cada região fisiográfica do Estado, a saber :

- Tipo Florestal AG 2, – Vegetação Arbustiva Arbórea Aberta, com espécies de altura média de 3,0 metros e emergentes de 5,0 metros; ocorrem em solos rasos arenosos e pedregosos, na maioria das vezes em áreas que indicam um processo de regeneração. Nestas áreas a topografia é suave e as maiores alturas são menores ou igual 550 metros.
- Tipo Florestal AG 3, – Vegetação Arbustiva Arbórea Fechada, caracterizada por espécies com altura média de 4,0 metros e indivíduos emergentes de 7,0 metros. Ocorre normalmente em áreas de topografia de suave a ondulada, em solos profundos do tipo areno-argilosos, sendo freqüente a presença da pedregosidade

e afloramentos rochosos. Este tipo de vegetação apresenta-se nas encostas e topos de serras entre as alturas de 550 a 750 metros.

- Tipo Florestal AG 4, – Vegetação Arbórea Fechada, com presença de alguns arbustos, altura média de 5,0 metros e emergentes com mais de 8,0 metros, ocorre principalmente nas encostas e nos topos de serras. Nestas áreas a topografia é de ondulada a fortemente ondulada e acontece nas alturas maiores de 750 metros, onde surgem os famosos “Brejos de Altitudes”.

2.2 MATERIAIS UTILIZADOS

Foram utilizados os seguintes materiais: carta topográfica, folha SC.24-X-B-III, escala 1:100.000, elaborada pela 3ª Divisão do Serviço Geográfico do Exército para a Superintendência de Desenvolvimento do Nordeste – SUDENE, em 1986; carta de vegetação em 1:100.000, elaborada em 1992, por meio de classificação visual de imagens do satélite Landsat TM-5, em papel fotográfico, composição colorida das bandas 3, 4 e 5; carta altimétrica, da área de estudo em formato digital dxf; imagem Landsat TM-5, 215/066, data de obtenção 07/05/2001, georeferenciada; GPS; máquina fotográfica digital; pentium II, e programas auxiliares.

2.3 METODOLOGIA

A metodologia utilizada neste trabalho desenvolveu-se no Sistema de Processamento de Informações Georreferenciadas-SPRING (Versão 3.6) e consistiu basicamente nas três etapas a seguir:

- Classificação da imagem que recobre a área de estudo pelo método supervisionado utilizando o algoritmo da máxima verossimilhança (MaxVer);
- Realização da integração dos dados auxiliares antes, durante e depois do processo de classificação da imagem;
- Analisar os resultados que forem gerados;

2.3.1 CLASSIFICAÇÃO DA IMAGEM DA A ÁREA DE ESTUDO

O Sistema SPRING permite a classificação supervisionada e não supervisionada. Portanto, nesta pesquisa, classificou-se a imagem, que recobre a área de estudo, pelo método supervisionado utilizando-se a técnica de máxima verossimilhança (MaxVer), com este sistema. O treinamento foi obtido com base no mapa de vegetação e na pesquisa de campo, foram utilizadas 45 amostras, sendo 15 amostras em cada classe.

Iniciando o processo de classificação, fez-se a composição das bandas 5, 4 e 3 da imagem e em seguida o realce. Para o treinamento supervisionado identificou-se no mapa de vegetação áreas representativas de cada classe e as correspondentes na imagem. Neste procedimento teve-se o cuidado de escolher amostras bastante homogêneas de cada classe na imagem, mas ao mesmo tempo procurou-se incluir toda a variabilidade dos níveis de cinza do tema em questão.

Este sistema permite também realizar, antes da classificação, a análise quantitativa da qualidade das amostras que participarão do treinamento no momento da classificação. Assim, foram eliminadas as amostras que apresentaram confusão média acima de 10%. De posse das amostras e das bandas escolhidas, a imagem foi classificada com um limiar de aceitação de 95%.

2.3.2 GERAÇÃO DOS DADOS AUXILIARES

A formação do modelo numérico de terreno-MNT foi iniciada com entrada da altimetria (formato (dxf)) no sistema por meio da importação dos dados. A partir desses dados gerou-se uma grade retangular de 30m x 30m (compatível com a resolução da imagem utilizada) e a declividade no modo MDT do sistema. O fatiamento foi gerado a partir da grade retangular e os temas da imagem temática tiveram como base nos intervalos de ocorrência dos tipos florestais (0 – 550, 550 – 750 e > 750 metros) identificados na pesquisa de campo e na observação dos mapas de vegetação e topográfico da área de estudo. Desta forma, um Plano de Informação de categoria numérica originou um Plano de Informação de categoria temática (fatiamento) representando um aspecto particular do modelo numérico de terreno. O sistema também permitiu a geração da imagem em níveis de cinza (de 0 a 255) para o modelo numérico de terreno, onde a resolução (em metros) da imagem de saída tem a mesma da grade retangular que a gerou.

2.3.3 INCORPORAÇÃO DOS DADOS AUXILIARES NO PROCESSO DE CLASSIFICAÇÃO ANTES, DURANTE E DEPOIS

Um Sistema de Informações Geográficas (SIG) é um ambiente computacional para integrar dados cartográficos, cadastrais, de sensores remotos redes e modelos numéricos de terreno. Permite também consultar, recuperar, visualizar, manipular e plotar o conteúdo de um banco de dados "georeferenciados". É uma ferramenta poderosa para análise e manipulação de um grande número de informações espaciais (Câmara, 1993).

A partir da interpretação de imagens orbitais de multisensores e multitemporais, são extraídas informações sobre as principais variáveis do meio físico (geomorfologia, hidrologia, solos, uso da terra e cobertura vegetal e clima) que influenciam nos fatores condicionantes do meio físico (altimetria, hipsometria,

declividade, exposição das vertentes e geomorfologia). Estas informações podem ser integradas e classificadas por meio de um SIG.

A metodologia da incorporação de dados auxiliares antes do processo de classificação de imagens é composta de quatro etapas: Em primeiro lugar gera-se três imagens binárias (BI(1), B(2) e BI(3)) fatiando-se o modelo numérico do terreno, onde considera-se os intervalos de alturas em relação a ocorrência dos tipos florestais. Sendo os seguintes intervalos: 0 – 550, 550 – 750 e > 750 metros. As regras para geração das imagens binárias são as seguintes:

- 1) Na geração de BI(1): MNT(1) → 0 para alturas menor ou igual a 750m e 1 para maior que 750m;
- 2) Na geração de BI(2): MNT(2) → 0 para alturas menor que 550m e maior que 750m e 1 para maior que 550m e menor ou igual que 750m;
- 3) Na geração de BI(3): MNT(3) → 0 para alturas maior que 550m e 1 para menor ou igual que 550m.

Fatia-se a imagem original por meio do cruzamento de cada banda original com a imagem binária correspondente ao intervalo de alturas onde ocorre cada tipo florestal, conforme o quadro 1. Assim, gera-se três composições de imagens fatiadas (TM543F(1), TM543F(2) e TM543F(3)) com um número menor de classes.

QUADRO 1 – CRUZAMENTO DAS BANDAS ORIGINAIS COM AS IMAGENS BINÁRIAS

Cruzamentos	Cruzamento	Cruzamento	Cruzamento
I	TM5 x B(1)	TM5 x B(2)	TM5 x B(3)
II	TM4 x B(1)	TM4 x B(2)	TM4 x B(3)
III	TM3 x B(1)	TM3 x B(2)	TM3 x B(3)
(I+II+III) Gera	TM543F(1)	TM543F(2)	TM543F(3)

Finalmente, faz-se o realce e a classificação separadamente de cada composição fatiada, juntando este conjunto chega-se a classificação final da imagem original com o uso de dados auxiliares antes do processo de classificação da imagem.

A metodologia da incorporação de dados auxiliares durante o processo de classificação de imagens consistiu na formação de uma composição das três bandas 5, 4 e 3 mais a imagem de níveis de cinza (altimetria) do MNT, que recobrem a área de estudo, gerando uma nova imagem e em seguida fazendo o realce e a classificação do conjunto pelo processo supervisionado utilizando-se o algoritmo da Máxima Verossimilhança-Maxver obtém-se a imagem classificada com o uso do dado auxiliar durante o processamento.

Incorporação de dados auxiliares após o processo de classificação de imagens. Neste caso classifica-se primeiro a imagem, que recobre a área de estudo, pelo

método supervisionado utilizando-se a técnica de máxima verossimilhança (MaxVer), e com os dados do MNT (fatiamento) faz-se a classificação dos pixels com base nos intervalos das alturas que definem a ocorrência das classes de tipos florestais. Este processo é realizado com o cruzamento de dois planos de informações PI1 (Imagem classificada) e PI2 (fatiamento). Para isto pode-se utilizar a programação LEGAL do SPRING ou outro sistema que permita o cruzamento de dois planos de informações.

3. CONSIDERAÇÕES FINAIS

No primeiro experimento, o mapa temático resultante da classificação da imagem em estudo com o algoritmo supervisionado MaxVer sem a utilização de dados auxiliares foi avaliado através da matriz de erro (contingência) das amostras de verificação, sendo verificado um índice de confusão média de 10.35%, Quadro 2.

QUADRO 2 - MATRIZ DE ERROS DA CLASSIFICAÇÃO TM 5, 4 e 3 (imagem 2001)

Classes temáticas	Verdade de campo				S. Linha
	Ag2	Ag3	Ag4	Abste	
Ag2	3597 19.09%	1006 5.34%	34 0.18%	209 0.11%	4846
Ag3	429 1.32%	1292 6.86%	77 0.41%	44 0.23%	1662
Ag4	26 0.14%	558 2.96%	11218 59.55%	529 2.81%	12331
Soma Col.	3872	2856	11329	782	18839

Classes temáticas	Exatidão do produtor	Exatidão do usuário
Ag2	92.90%	77.57%
Ag3	45.24%	79.85%
Ag4	99.02%	95.05%

Colunas : dados de referencia
 Desempenho geral: 85.50 %
 Confusão média : 10.35 %
 Abstenção media : 4.15 %
 Estilística KHAT : 74.09 %
 Variância KHAT : 1.609e-005

Estilística TAU : 78.25 %

Num segundo experimento, efetuou-se a classificação utilizando todas as variáveis, isto é, a imagem multiespectral e os dados de altitude, representados como uma quarta banda. O resultado obtido revela que a confusão média entre classes diminuiu significativamente de (1.11%), principalmente no topo dos morros, onde a vegetação do tipo arbóreo é mais frequente, Quadro 3.

QUADRO 3 - MATRIZ DE ERROS DA CLASSIFICAÇÃO DA COMPOSIÇÃO NC(altimetria),5, 4 e 3 (imagem) 2001

Classes temáticas	Verdade de campo				S. Linha
	Ag2	Ag3	Ag4	Abste	
Ag2	2319 0.50%	47 0.11%	0.00 0.00%	161 0.38%	2527
Ag3	421 1.00%	7698 18.26%	0.00 0.00%	404 0.96%	8523
Ag4	0.00 0.00%	0.00 0.00%	28968 68.72%	2138 5.07%	31106
Soma Col.	2740	7745	28968	2703	42156

Classes temáticas	Exatidão do produtor	Exatidão do usuário
Ag2	84.64%	98.01%
Ag3	99.39%	94.81%
Ag4	100.00%	100.00%

Colunas : dados de referencia

Desempenho geral: 92.48 %

Confusão média : 1.11 %

Abstenção média : 6.41 %

Estatística KHAT: 84.61 %

Variância KHAT: 5.297e-006

Estatística TAU : 88.72 %

A incorporação de dados auxiliares no processo de classificação do sensoriamento remoto é uma alternativa importante em estudos baseados especificamente na análise per-pixel da informação. Entretanto a escolha de variáveis a serem incluídas é crítica. O bom senso sugere que analistas devem escolher inteligentemente, somente variáveis de significado prático e conceitual para

o problema da classificação que se apresenta. Incorporar dados auxiliares ilógicos ou suspeitos podem consumir rapidamente recursos limitados para a análise de dados e conduzir a resultados inaccurados (Jensen, 1986).

Neste trabalho apresentou-se a metodologia para dados auxiliares. A implementação do uso destes dados está em fase de desenvolvimento.

4. REFERÊNCIAS BIBLIOGRÁFICAS

CÂMARA, G. Anatomia de sistemas de informações geográficas: visão atual e perspectivas de evolução. In: ASSAD, E.D., SANO, E.E. Sistema de informações geográficas: aplicações na agricultura. Brasília, EMBRAPA, 1993.

CROSTA, A. P. Processamento digital de imagens de sensoriamento remoto. Campinas: IG/UNICAMP, 1992.

HUTCHINSON, F. C. Techniques for Combining Landsat and Ancillary Data for Digital Classification Improvement. **Photogrammetric Engineering and Remote Sensing**, 48(1):123-130, 1982.

JENSEN, J. R. Introductory Digital Image Processing: A Remote Sensing Perspective. 2nd. Ed., Saddle River, N. J. : Prentice-Hall, 1986, 316p.

SILVA FILHO, A. A. et al. Mapeamento do Cobertura Florestal Nativa Lenhosa do Estado de Pernambuco. In: **Desenvolvimento Florestal Integrado no Nordeste.** FO: BRA/87/007, Documento de Campo N^o17, Recife, 26p, 1998.

STRAHLER, A. H. et al. Improving forest Cover Classification accuracy from Landsat by incorporation topographic information. **Proceedings of the Twelfth international Symposium on remote sensing of environment**, Michigan, p. 927-942, 1978.

5. AGRADECIMENTOS

Agradecemos ao departamento de engenharia cartográfica – Decart da Universidade Federal de Pernambuco por disponibilizar sua estrutura.

A secretária de Ciência, Tecnologia e Meio Ambiente por disponibilizar sua estrutura e os materiais.