

ISSN 15169111

PAPERS DO NAEA Nº 329

**QUALIFICANDO PIXELS, SIGNIFICANDO MAPAS:
NOTAS SOBRE RECLASSIFICAÇÃO DE IMAGENS
DIGITAIS SOB O TESTEMUNHO DE ESTATÍSTICAS E
AVALIAÇÃO DA ENTROPIA DA INFORMAÇÃO**

Francisco de Assis Costa

Belém, Setembro de 2014

O Núcleo de Altos Estudos Amazônicos (NAEA) é uma das unidades acadêmicas da Universidade Federal do Pará (UFPA). Fundado em 1973, com sede em Belém, Pará, Brasil, o NAEA tem como objetivos fundamentais o ensino em nível de pós-graduação, visando em particular a identificação, a descrição, a análise, a interpretação e o auxílio na solução dos problemas regionais amazônicos; a pesquisa em assuntos de natureza socioeconômica relacionados com a região; a intervenção na realidade amazônica, por meio de programas e projetos de extensão universitária; e a difusão de informação, por meio da elaboração, do processamento e da divulgação dos conhecimentos científicos e técnicos disponíveis sobre a região. O NAEA desenvolve trabalhos priorizando a interação entre o ensino, a pesquisa e a extensão.

Com uma proposta interdisciplinar, o NAEA realiza seus cursos de acordo com uma metodologia que abrange a observação dos processos sociais, numa perspectiva voltada à sustentabilidade e ao desenvolvimento regional na Amazônia.

A proposta da interdisciplinaridade também permite que os pesquisadores prestem consultorias a órgãos do Estado e a entidades da sociedade civil, sobre temas de maior complexidade, mas que são amplamente discutidos no âmbito da academia.

Papers do NAEA - Papers do NAEA - Com o objetivo de divulgar de forma mais rápida o produto das pesquisas realizadas no Núcleo de Altos Estudos Amazônicos (NAEA) e também os estudos oriundos de parcerias institucionais nacionais e internacionais, os Papers do NAEA publicam textos de professores, alunos, pesquisadores associados ao Núcleo e convidados para submetê-los a uma discussão ampliada e que possibilite aos autores um contato maior com a comunidade acadêmica.



Universidade Federal do Pará

Reitor

Carlos Edilson de Almeida Maneschy

Vice-reitor

Horacio Schneider

Pró-reitor de Pesquisa e Pós-graduação

Emmanuel Zagury Tourinho

Núcleo de Altos Estudos Amazônicos

Diretor

Durbens Martins Nascimento

Diretor Adjunto

Ana Paula Vidal Bastos

Coordenador de Comunicação e Difusão Científica

Silvio Lima Figueiredo

Conselho editorial do NAEA

Profa. Dra. Ana Paula Vidal Bastos

Prof. Dr. Armin Mathis – NAEA/UFPA

Prof. Dr. Durbens Martins Nascimento – NAEA/UFPA

Profa. Dra. Edna Castro – NAEA/UFPA

Prof. Dr. Fábio Carlos da Silva – NAEA/UFPA

Prof. Dr. Francisco Costa – NAEA/UFPA

Prof. Dr. Luis Eduardo Aragón Vaca – NAEA/UFPA

Prof. Dr. Silvio Lima Figueiredo – NAEA/UFPA

Sector de Editoração

E-mail: editora_anae@ufpa.br

Papers do NAEA: papers_anae@ufpa.br

Telefone: (91) 3201-8521

Paper 329

Recebido em: 10/08/2014.

Aceito para publicação: 09/09/2014.

Revisão de Língua Portuguesa de responsabilidade do autor.

QUALIFICANDO PIXELS, SIGNIFICANDO MAPAS: NOTAS SOBRE RECLASSIFICAÇÃO DE IMAGENS DIGITAIS SOB O TESTEMUNHO DE ESTATÍSTICAS E AVALIAÇÃO DA ENTROPIA DA INFORMAÇÃO

Francisco de Assis Costa¹

Resumo:

O artigo objetiva indicar caminhos metodológicos para que as estatísticas disponíveis sejam usadas para a significação dos mapas digitais, de modo que estes representem melhor o conhecimento sobre estruturas e dinâmica econômicas e sociais. Sugerindo o uso dos recursos de representação computacional do espaço por uma perspectiva que considera o papel decisivo da informação em contexto de *ação comunicativa*, indica o uso da *entropia da informação* como métrica de qualificação dos mapas, em complemento ao *índice kappa*.

Palavras-Chave: Mapas digitais. Qualificação de imagens. Significação de imagens.

Mapas digitais. Entropia da informação.

Abstract:

The article aims to indicate methodological ways to use the available statistics to signifying of digital maps, so that they better represent knowledge about economic and social structures and their dynamics. Treating the use of computational resources to representation of space in a perspective that considers the critical role of information in the context of communicative action, the article indicates the use of "information entropy" as a metric for qualification of maps, in addition to the "kappa index".

Keywords: Digital maps. Qualification of images. Meaning of images. Digital maps.

Information entropy.

¹ Professor Associado do NAEA e do Programa de Pós-Graduação em Economia da UFPa. Pesquisador Associado da RedeSist (UFRJ) e da Rede Geoma (O distanciamento entre realidades sociais e expressões digitais da MCT).

1. NOTAS INTRODUTÓRIAS: O DISTANCIAMENTO ENTRE REALIDADES SOCIAIS E EXPRESSÕES DIGITAIS DA AMAZÔNIA RURAL

A dinâmica rural na Amazônia tem sido observada privilegiando o fenômeno do desmatamento (COSTA, 2009). A leitura do processo feita pelo principal instrumento de rastreamento por imagem de satélite, o Projeto de Estimativas de Desflorestamento da Amazônia (PRODES), do Instituto Nacional de Pesquisas Espaciais (INPE), é binária: trata-se de acompanhar por sensoriamento remoto a “Taxa de Desflorestamento Bruto”, representado pelas áreas desmatadas em corte raso em dado ano, as quais se somam ao “passivo ambiental” do ano anterior compondo o novo saldo. Assim, favorecendo um viés ambientalista estreito, para o qual o desmatamento é um mal em si, a metodologia silencia sobre razões e sujeitos, bem como ignora contramarchas: dinâmicas, protagonizadas por agentes e atores diversos em processos decisórios complexos, que se fazem eventual ou sistematicamente na direção contrária ao desmatamento.

A metodologia do PRODES contraria, há tempos, tanto informações primárias de campo, resultantes de pesquisas antropológicas, sociológicas e econômicas, quanto informações secundárias, resultantes dos grandes levantamentos estatísticos do Instituto Brasileiro de Geografia e Estatísticas (IBGE) na Amazônia. Ainda nos anos noventa, pesquisadores da Universidade de Indiana indicavam a necessidade de se instruir, com resultados de pesquisa *in loco*, em correspondência espacial estrita, os algoritmos de leitura de imagem com os fundamentos da ação humana que informam sobre dinâmicas regenerativas, de restauração da cobertura vegetal na Região (MORAN, PACKER, BRONDIZIO, TUCKER, 1996; MCCRACKEN, BRONDIZIO, NELSON, MORAN, SIQUEIRA E RODRIGUEZ-PEDROZA, 1999). Ao longo dos anos de 2000, nossas próprias pesquisas demonstravam serem as estatísticas do IBGE capazes de demonstrar os volumes e localizações, mesmo que não georeferenciadas, dessas áreas em regeneração em toda a Amazônia, seja em forma de mata secundária, seja em forma de plantios de culturas permanentes ou de silviculturas, nos múltiplos sistemas produtivos que compunham o agrário regional (COSTA, 2001; COSTA, 2006; COSTA, 2009). Os resultados expressavam um balanço abrangente, se afigurando extraordinariamente relevantes, eis que, apenas as áreas de capoeiras representavam 4,5 milhões de hectares em 1995 – em torno de 10% de toda a área apropriada na região naquele ano...

O longo esforço de pesquisa do grupo de Indiana, já mencionado, se fez com o propósito de criar validação de campo para a constituição de assinaturas espectrais que correspondessem à complexidade dos sistemas de produção em desenvolvimento na região (MORÁN, BRONDIZIO, 1994). A busca de assinaturas espectrais tem sido um modo eficaz de aplicar as técnicas de sensoriamento remoto no planejamento e acompanhamento das atividades rurais em todo o mundo. Há que considerar, todavia, que tal eficácia, observada pela relação custo/benefício, se move na razão inversa da regularidade de forma e conteúdo dos sistemas de produção rural em consideração e tratamento. Quanto mais homogêneos e regulares os plantios, menor esforço de pesquisa será necessário para validação, menor custo de produção da assinatura, menor custo de sua aplicação por unidade de área reconhecida; quanto mais diversos os sistemas, e mais irregulares as formas de ocorrência, maior o custo de produção de assinatura e maior o custo de reconhecimento.

Na Amazônia, demonstramos em pesquisas recentes (COSTA, 2009B; COSTA, 2012; COSTA, 2013) que o rural se estrutura a partir da dinâmica de trajetórias tecnológicas bem distintas, se destacando padrões nos quais as trajetórias patronais apresentam graus de especialização elevados, enquanto as camponesas se caracterizam pela diversidade, mesmo quando certas atividades-chaves ganham proeminência. Ademais, as formas como os sistemas se configuram, em uma única trajetória camponesa, são extraordinariamente diversas, em territórios distintos, encarecendo significativamente para essa forma de produção e suas trajetórias as possibilidades da produção de assinaturas espectrais válidas no tempo e no espaço, baseadas em reconhecimento empírico, físico, *in loco*. Mediante tais restrições, corroboram-se, com os poderosos instrumentos de leitura por imagem georeferenciada, importantes assimetrias de informação sobre a realidade rural da região: as estruturas produtivas patronais, de leitura mais barata e simples, ganham expressão e nitidez, enquanto as camponesas, de leitura complexa e custosa, se obscurecem em expressões borradas e confusas. Assimetrias de informação não são neutras, confirmando assimetrias de outras ordens, econômico-sociais e ecológicas.

Acresce, ampliando o viés e seu impacto, o desequilíbrio na expertise dos órgãos de governo e de seus técnicos, no que se refere às diferentes estruturas, seus gestores e expressões territoriais. Exemplo a considerar de tal desequilíbrio de signos e significados transportado para a leitura da realidade, a plataforma TerraClass. Um avanço inegável por

reconhecer as insuficiências do PRODES, já apontadas, a mencionada plataforma lançada em 2011 pelo INPE explora as faixas intermediárias de espectros das imagens e convoca o conhecimento tácito da Empresa Brasileira de Pesquisa Agropecuária (Embrapa) na condição de “testemunha” e “verdade de campo”. Os resultados realçam os grandes tratos de “culturas anuais” e de “pasto”, por inferência de tamanho e conteúdo, patronais, em múltiplos formatos, “limpos”, “sujos”, “em regeneração”, deixando, literalmente, grandes áreas cinzentas de “vegetação secundária” e de “mosaicos de ocupação” indistintos, por inferência, camponeses. Novamente, grande foco de luz, lá, sombras e borrões, aqui...

Realidades que não são vistas, ou são observadas deformadas ou fragmentadas terão grande probabilidade de, ou não serem tratadas pela política, ou abordadas com equivalentes distorções. Há que se dispor de meios de compensar as obscuridades, de aclará-las, de modo a minimizar o risco de injustiças derivadas de uma “retórica de mapas...” (ALMEIDA, 1994).

Nesse aspecto pretende contribuir o presente artigo, primeiro, indicando caminhos para que o uso das estatísticas disponíveis seja base de significação dos *mapas de pixels* em circulação – e estas bases de representação do conhecimento que, em expressões estatísticas, revelam o mundo da vida: a interação entre as bases produtivas territorialmente ancoradas e os fluxos de seus resultados. Segundo, sugerindo o uso dos recursos de representação computacional do espaço por uma perspectiva da informação em contexto de ação comunicativa.

2. PONTOS E PIXELS

Pontos são constitutivos de última instância das imagens, matéria prima das retas e curvas, dos segmentos e manchas que se combinam em representações gráficas do mundo. É enunciado clássico da geometria euclidiana – esta maneira de representação do mundo por combinação de suas formas elementares, ponto, reta e plano –, que já no menor segmento de reta, em qualquer dos traços que formam as imagens, os pontos ocorrem em seqüências infinitas.

Até o surgimento das imagens digitalizadas, as representações gráficas se fazem apenas como morfologia mimética, como analogia, portanto, seja por imitação de expressões fenomênicas percebidas e interpretadas por agentes, que controlam, por habilidade mental e

corporal, a justaposição de traços e manchas, claros e escuros, luzes e sombras, seja por captura direta, sem mediação cognitiva, do fenômeno de refletância da luz sobre o objeto representado sobre material sensível. Pontos, em qualquer desses casos, são abstrações reais – eles existem, é certo, porém, numa dimensão perceptível apenas por elaboração cognitiva.

Nas imagens óticas digitalizadas, os pontos, em sua realidade essencial, em sua condição de partícula elementar do espaço, ente geométrico de dimensão, comprimento, largura e altura iguais ou tendentes a zero, ocorrem em “pacotes”, nos *pixels*, semelhantemente ao que se verifica nos “quantas” do eletromagnetismo. Nas imagens óticas digitalizadas de porções da superfície terrestre, os pontos que compõem coletivamente um pixel são pontos geodésicos, é dizer, pontos materializados no terreno, onde as coordenadas de sua posição real podem ser relacionadas com informações contidas na especificidade da imagem e com a especificidade do terreno. De modo que, nas imagens digitalizadas da terra, os magotes de pontos encapsulados em *pixels* informam sobre a especificidade de partes elementares do objeto. Quanto menor o pixel, menor a porção de pontos (que na noção euclidiana continua infinita), ou, mantido o número, menores os pontos (de tamanhos que se aproximam cada vez mais de zero) reais nele abrigados, maior, portanto, por um movimento ou por outro, o poder de expressão das respectivas especificidades.

Como portadores de informações sobre pontos com especificidades, os pixels tornaram-se signos de linguagens capazes de descrever realidades com a virtuosidade das capacidades e desenvolvimentos computacionais. Todavia, de imediato, são ajuntamentos de sinais sem significado para além dos termos de sua manifestação: nenhuma ou máxima refletância (completa escuridão ou total claridade), brilho intenso ou esmaecido, refletância e brilho intermediários: muitíssimos tons de luz e sombra, equivalentes tons de cinza.

3. PIXELS E MAPAS

Uma fotografia digital de uma área, com as referências geográficas que associam pontos no espaço real com pixels da imagem, constitui um *mapa de pixels*.

Cada pixel pode se apresenta em 2^n tons de cinza, a depender de n , o número de bits que define a resolução radiométrica do sensor. Alguns sensores, como o IKONOS, possuem de 16 bits, permitindo 65.536 (2^{16}) níveis de cinza. A depender do “ponto de vista” (distância

e posição relativa do observador) e das habilidades do aparato de leitura (número de pixels abarcado pela “retina”), o pixel pode representar diferentes unidades espaciais: de 30 m x 30 m nas imagens de satélite LandSat a apenas 60 cm x 60 cm em imagens QuickBird.

Um conjunto de pixels vizinhos, conformando uma mancha completamente escura; uma linha de pixels claros desenhando um segmento de reta ou uma curva, são figuras sensorialmente e intuitivamente identificáveis, sobre as quais se poderá, até, enunciar por analogia. Mas não é se prestando a analogias que o mapa de pixels demonstra sua força de expressão. Sua maior capacidade está em se deixar identificar e controlar no nível mais elementar e, assim, se deixar processar por técnicas estatísticas e computacionais (análises de *clustering*, modelos baseados em autômatos computacionais, etc.) baseados em valores numéricos de atributos fundamentais dos pixels, como os da refletância, os quais, combinadas com análises subseqüentes ou concomitantes de forma (geometria dos aglomerados) e conteúdo (pigmentação, granulometria), os classificam com extraordinário detalhamento e precisão. Delineados os aglomerados, facultam-se exercícios que os relacionam por comparação, padronização, justaposição ou confrontação. De tudo, emerge uma semântica instrumental, uma linguagem, uma codificação de sinais, que, *nos seus próprios termos*, é capaz de descrever o mundo em mapas com enorme riqueza de detalhes, com uma multidão de objetos, tanto mais, a depender das regras de *clustering*, dos mais ou menos estreitos limites de padronização e segregação que os orientam.

Tais objetos, as “coisas” e suas relações designadas nessa linguagem seminal, contudo, nada mais que nuvens de pixels sob absoluto controle numérico e estatístico, não se revelam de imediato para a realidade social; eles carecem de significação, isto é, de atribuição de sentidos por interpretações à luz de teorias e experiências. Bem diferentes, pois, nesse aspecto, os *mapas de pixels* dos mapas analógicos construídos a partir de significados *a priori*, dos quais procuram ser evidentes portadores.

4. MAPAS E INFORMAÇÃO

Significação de pixels ou classes de pixels implica em interpretação – passo fundamental no processo de comunicação. Por seu turno, um *mapa de pixels* é parte de um processo de comunicação - ou não é nada.

Um processo genérico de comunicação envolve interação entre agentes, recursos e métodos na produção, envio e absorção de informações: sinais codificados pelos agentes emissores em linguagem compreensível pelos agentes receptores em uma interação comunicativa. Para as mesmas informações ou mensagens se poderão ter diferentes significações, a depender do sistema conceitual (ou físico) de referência (SHANNON, WEAVER, 1963). As referências conceituais são objetos de acordo entre os agentes porque delas dependem maior ou menor poder de equacionamento dos respectivos problemas e, portanto, maior ou menor capacidade de enunciado e elucidação das imagens. Assim, nos termos de (inter)ação comunicativa (HABERMAS, 1995), na qual se posicionam interlocutores com interesses estratégicos relativizados por conhecimentos se *realizam* as informações, i.e. elas cumprem seu papel como “... redução de incerteza oferecida quando se obtém respostas a uma pergunta” (EPSTEIN, 1986:35). Precisamente aqui, ao cumprir seu papel de última instância, as mensagens poderão ser objetivamente avaliadas (THEIL, 1967).

Mapas de pixels constituem *recursos* de processos de comunicação: eles são, em cada caso, o canal (a *media*) e a codificação básica. A transformação das informações de um *mapa de pixels* em *mapa de realidade* que oriente a resolução de problemas social ou privadamente acordados exige métodos de transmutação da *codificação básica* em *símbolos carregados de significados*. O maior ou menor sucesso nessa passagem definirá a eficiência do *mapa de realidade* e, por conseguinte, do método que o organizou. Como medida dessa eficiência a *entropia da informação* – um *mapa de realidade* (um *mapa de pixels* transmutado por significação) se sobressairá diante de outros na medida em que a *entropia da informação total* que produza em relação à realidade representada seja maior que as demais.

Proposta por Shanon (1948), e encampada por Theil (1967), a *entropia da informação* se baseia na idéia de que a quantidade de informação contida em uma mensagem, a sua entropia (H_i), depende diretamente da incerteza que reina em torno do evento informado (E_i) e inversamente com a probabilidade de sua ocorrência (p_i). Se o evento E_i ocorre, a carga de informação que a mensagem que dele deu conhecimento será tanto maior quanto menor seja p_i . Se é sempre o evento i que ocorre ($p_i = 1$) então a carga de informação que seu comunicado porta é zero. Se um evento ocorre muito raramente, ($p_i \rightarrow 0$), e, por isso, há muita incerteza em torno dele, a quantidade de informação de sua ocorrência é, relativamente, muito elevada.

A equação (1) expressa essas propriedades da informação e, no seu enunciado geral, garante que, reconhecida a possibilidade de que um sistema se explicita por completo por n fluxos de mensagens, a quantidade de informação total que produzirá (H , expresso em bits, dada a base 2 do logaritmo; se o logaritmo é natural, então *nats*) será tanto maior quanto melhor distribuídas forem as probabilidades de realização respectivas.

$$H = - \sum_{i=1}^n p_i \log_2 p_i \quad (1)$$

Assegura, ademais, a relação (1), que a ocorrência de eventos de probabilidades menores produza efeitos na entropia da informação (na redução de incerteza) mais que proporcionais aos que simetricamente se realizam com eventos de probabilidades maiores; que quanto maior o grau de concentração da informação, menor a entropia, até o ponto em que, se apenas uma informação tem probabilidade de se apresentar, então a entropia é zero; que, afinal, se houver equidistribuição das probabilidades dos fluxos de informação, a entropia é máxima ($p_i=1/n$). Assim, a entropia máxima dos sistemas de informação cresce com n , com o número de fluxos de mensagens. Em perspectiva evolucionária, se assevera com isso que um sistema amplia sua capacidade de produzir informação com diversificação e complexificação (FENZL, HOFKIRCHNERB, s. d.).

A relação entre a entropia total (a carga real de informações que um sistema produz) e a entropia máxima (a carga limite de informações dado número de canais do sistema) produz a entropia relativa da informação: a proporção (h , de valor entre 0 e 1) realizada da capacidade informacional total do sistema. Como segue:

$$h = \frac{H}{H_{max}} \quad (2)$$

5. INFORMAÇÃO E VERDADE

Em que medida, porém, o pixel (re)significado (ou com significado) corresponde à realidade concreta? Em que medida subjetividade e ambigüidade dos agentes ou as características ontológicas das teorias envolvidas na classificação (SOWA, 2006) ou, ainda e

simplesmente, as formas de agregação e as técnicas de leitura dos elementos classificados criaram vieses e distorções (DIAS et al, 2002)?

(Re)significado, o pixel precisa ser julgado, à luz de sua aderência ao mundo da vida. O recurso mais estabelecido é o conhecido como julgamento por múltiplos juízes. Proposto por Cohen (COHEN, 1960), o Coeficiente Kappa é uma forma de estabelecer graus de concordância inter-juízes (ou inter-amostal) na validação de um processo de significação. Considerando que as unidades em análise são independentes, que as categorias da classificação são exaustivas, exclusivas e, também, independentes, assim como independentemente agem os juízes (estatisticamente, a covariância entre as amostras é zero) o coeficiente k , uma grandeza que varia entre 0 e 1, pode ser definido como a proporção de acordo entre os juízes (P_0) após ser retirada a proporção de acordo devido ao acaso (P_a), como na fórmula (2):

$$k = \frac{P_0 - P_a}{1 - P_a} \quad (3)$$

Numa outra perspectiva, k pode se entendido como o grau de acerto de uma regra, ou teoria, de significação de um mapa de pixels.

6. CONTROLANDO OS PIXELS DOS MAPAS

Uma delimitação espacial real A_R , da qual se conhecem os limites, composta por z_R pontos de coordenadas geográficas x,y , fotografada por equipamento digital resultou em imagem formada por um conjunto A_F de z_F pixels a_{ij} . Uma primeira descrição de A_R com base em A_F implicaria no conjunto A_α :

$$A_\alpha = \left\{ a_{ij} \mid (a_{ij} = l^2) \in A_F \wedge (i = 1, \dots, I; j = 1, \dots, J) : \Leftrightarrow (x, y) \in A_R \right\} \quad (4)$$

A_α é composto por $n_{A_\alpha} = I.J$ pixels $a_{ij} = l^2$, em que i, j , equivalente lógico de qualquer x,y no plano real correspondente, estabelece a posição do “ponto” de refletância e brilho

homogêneos e forma regular que representa a área de um quadrado no espaço real com l de lado. A área de A_R que corresponde à descrição feita em $A\alpha$ comporá um conjunto unitário:

$$\hat{A}_R = \left\{ a_R \mid a_R = \sum_{i,j=1}^{l,l} a_{ij} = n_{A_\alpha} \cdot l^2 ; n_{\hat{A}_R} = 1 \right\} \quad (5)$$

Por seu turno, a cada pixel a_{ij} correspondem atributos: refletância, forma, diversidade etc. Os elementos de $A\alpha$ podem ser distribuídos em diferentes subconjuntos em função de regras escalares aplicadas a cada tipo de atributos. Por exemplo, em função dos valores de refletância, contidos no conjunto:

$$A_{\alpha r} = \left\{ r_{ij} \mid r_{ij} = \text{refletância de } a_{ij} \in A_\alpha \right\} \quad (6)$$

Uma regra binária de formação de classes por escala regular estabeleceria duas faixas de ocorrência: uma entre $r = 0$ e $r = \{[1/2 \cdot \max(r)] - 1\}$ e outra entre $r = [1/2 \cdot \max(r)]$ e $r = \{[2/2 \cdot \max(r)] - 1\}$. Se a regra é que se verifiquem os padrões em três faixas de refletância distribuídas em escala regular, então a primeira faixa seria $r = 0$ e $r = \{[1/3 \cdot \max(r)] - 1\}$, outra entre $r = [1/3 \cdot \max(r)]$ e $r = \{[2/3 \cdot \max(r)] - 1\}$ e, por fim, outra entre $r = [1/3 \cdot \max(r)]$ e $r = \{[2/3 \cdot \max(r)] - 1\}$.

Generalizando: estabelecidas classes $s = 1, \dots, S$ na faixa de valores de refletância entre $\{[(s-1)/S] \cdot \max(r)\}$ e $\{[s/S] \cdot \max(r)] - 1\}$, se formarão conjuntos A_s contendo os pixels a_{ij} das classes respectivas, como segue:

$$A_s = \left\{ a_{ij} \mid a_{ij} \in A_\alpha \wedge \left[\frac{s-1}{S} \cdot \max(r_{ij}) \right] < r_{ij} < \left[\frac{s}{S} \cdot \max(r_{ij}) \right], r_{ij} \in A_{\alpha r}, \forall s = 1, \dots, S \right\} \quad (7)$$

Os elementos de A_s ocorrem com frequência n_{A_s} , ao tempo que a um conjunto composto pelos S subconjuntos A_s , corresponde um conjunto \hat{A}_S com S elementos equivalentes às áreas de cada subconjunto A_s . Assim,

$$A_S = \bigcup_{s=1}^S A_s \leftrightarrow \hat{A}_S = \left\{ a_s \mid a_s = n_{A_s} \cdot l^2 ; \forall s = 1, \dots, S \right\} \quad (8)$$

O espaço total de A_R , além do somatório de pixels em A_α , pode resultar da agregação por classes, de tal modo que:

$$\hat{A}_R = \left\{ a_R \mid a_R = \sum_{i,j=1}^{I,J} a_{ij} = \sum_{s=1}^S a_s = n_{A_\alpha} \cdot l^2 ; n_{\hat{A}_R} = 1 \right\} \quad (9)$$

6.1. CLASSES E OBJETOS DE PIXELS

Os conjuntos A_S agregam pixels em classes. Pixels da mesma classe, todavia, não ocorrem como *continuum* em A_α , dado que as classes são fenômenos espaciais descontínuos em A_r . As diferenças nas classes e as diferenças no interior das classes, particularmente as distinções espaciais, mas também as de formas, configuram o *mapa de pixels*. Controlam-se as discontinuidades espaciais de pixels da mesma classe estabelecendo a contigüidade como critério de formação de conjuntos de pixels A_{sc} , como indicado abaixo.

Considere-se, como acima, o conjunto de pixels que pertencem a uma classe s agrupada num subconjunto A_s . Este conjunto, em qualquer circunstância, seria a união dos conjuntos A_{sc} nele contidos com seus complementos, conforme estabelecido em (10). No momento em que não há nenhum A_{sc} , A_{sc} é um conjunto vazio e o complemento de A_{sc} é A_s . Se se toma o elemento (pixel) de menores coordenadas em A_s (conf. (11)) e se verificam os elementos em torno dele em cadeia ininterrupta de vizinhança, é dizer, até que ocorra um $a_{ij} \notin A_s$ (conf. (12)), dispõe-se dos elementos que formam um primeiro A_{sc} . O complemento de A_{sc} , agora, de onde se elegerá o elemento de menores coordenadas, conterá apenas os elementos de A_s que não compuseram o objeto descrito pelos elementos de A_{sc} . Assim, se formarão sucessivos subconjuntos A_{sc} até que o complemento de A_{sc} seja um conjunto vazio.

$$A_s = A_{sc} \cup \overline{A_{sc}} \quad (10)$$

$$A_T = \{ a_{kp} \mid a_{kp} = a_{ij} \in \overline{A_{sc}} \wedge [k = \min(i) ; p = \min(j) ; n_{A_T} = 1] \} \quad (11)$$

$$A_{sc} = A_T \bigcup_{\substack{i=k+t \\ j=p+t}}^{a_{ij} \in A_s} A_t = \{a_{ij} \mid a_{ij} \in A_s\} \quad (12)$$

$$A_s = \bigcup_{c=1}^C A_{sc} \leftrightarrow \dot{A}_s = \{a_{sc} \mid a_{sc} = n_{A_{sc}} \cdot l^2; \forall c = 1, \dots, C\} \quad (13)$$

Ao final, têm-se os seguintes balanços: o espaço total de A_R , além das formas de agregação apresentadas em (9), poderá ser o somatório dos objetos contidos em cada classe de pixels. De modo que:

$$\dot{A}_R = \left\{ a_R \mid a_R = \sum_{i,j=1}^{I,J} a_{ij} = \sum_{s=1}^S a_s = \sum_{s,c=1}^{S,C} a_{sc} = n_{A_R} \cdot l^2; n_{A_R} = 1 \right\} \quad (14)$$

6.2. CLASSES DE PIXELS, VERDADE E INFORMAÇÃO

Até aqui, estruturou-se um conjunto de signos codificados em classes e essas codificadas em objetos. Obedecendo à mesma lógica o processo de codificação poderá seguir, qualificando os objetos pela consideração, um a um, dos diversos atributos dos pixels. Os resultados serão, em qualquer dos casos, passíveis de descrição como conjuntos A_s – cada um deles coleção de A_s , por sua vez coleção de A_{sc} e assim por diante – as quais conformam *mapas de pixels* em diversas formas e composição.

O que comunicam esses mapas depende do que significam as classes. As classes se construíram até este ponto com *critérios dos pixels*. O que elas abrigam sob *critérios de realidade* depende do ponto de vista, isto é, da estrutura conceitual que decodifica a realidade, especifica a partir daí seus problemas e elege o rol de eventos relevantes que poderiam estar abrigados nas classes de pixels. Digamos que a classe s abriga, sem saber, M aspectos m da vida real, entendidos como fenômenos relevantes por uma teoria considerada razoável pelos

agentes envolvidos². Mediante o fato de que o universo dos pixels pouco sabe sobre os m possíveis na realidade, apenas que, se eventos mutuamente exclusivos, qualquer um deles pode ocorrer com exclusividade e probabilidade 1 todos podem ocorrer em conjunto com probabilidades desconhecidas cuja soma é igual a 1, então s só pode se associar a m como classe genérica, tal que s será m_1 , ou m_2 , etc., ou, por fim, m_M . Assim, ter-se-á no *mapa de pixels* um conjunto de recodificações como segue:

$$C_s = \{s \mid s = \sum_{i=1}^M m_i ; n_{C_s} = 1; p(m_i) = 0; p(s) = 1; \} \quad (15)$$

Processos como esses, de recodificação generalizante, por ignorância, tem implicações que convém aludir. Por um lado, habilita o *mapa de pixels* a um teor de verdade formal máximo. Aclaremos esse ponto: se numa classe de refletância sabe-se, por ciência ou experiência, que se abrigam todas as “culturas temporárias”, as “pastagens plantadas” e as “pastagens naturais” de um A_R , mas não se conhece com que proporções esses eventos realmente ocorrem, os três fenômenos m serão fundidos em uma única designação genérica, produzindo uma recodificação de s que pode ser, eventualmente, “agropecuária” (= “culturas temporárias” ou “pastagens plantadas” ou “pastagens naturais”) que com probabilidade 1 se verificará na realidade. Se isso ocorre, uma medida do acerto da recodificação que relacione o significado imputado, com realidade encontrada, como o mencionado Coeficiente Kappa (k), tende a ser máximo, porque, mesmo que um juiz diga que o que encontrou na realidade foram “culturas temporárias”, outro que se trata de “pastagens plantadas” e um terceiro que se trata de “pastagens naturais”, todos concordam que se trata de “agropecuária”.

Por outro lado, a entropia da informação, em resposta, terá seu máximo contido pelo número de classes, $1/S$. De modo que o grau de verdade que k expressa para o mapa em questão é contra-arrestado pela redução de sua eficácia comunicativa, H .

7. ESTATÍSTICAS E INFORMAÇÃO

² Por exemplo, nossos estudos das trajetórias tecnológicas agrárias na Amazônia indicam que os tipos de produtos e suas combinações constituem uma fenomenologia importante para entender evolução e tendências das trajetórias. Nesse caso, m constituiria todos os produtos cuja base botânica tivesse refletância compatível com a classe s .

Entendemos, aqui, como *informações estatísticas* aquelas que se referem a uma delimitação espacial A_R com mais de dois pontos ($n_{AR} > 1$) sem se referirem aos seus pontos.

Trata-se de situação bem diferente da anterior, em que A_α é uma coleção de signos que se referem diretamente aos pontos e, assim, detalham o espaço real, encontrando nele, além do mais, um método descritivo dos seus elementos constitutivos (aglomerados mais escuros ou mais claros, de formas arredadas ou retilíneas, etc.) e explanatório de suas relações espaciais (contigüidade, afastamento, etc.). Quando se trata de informações estatísticas temos conjuntos de informações B (informações sobre tipos de uso, informações sobre dimensões, informações sobre resultados, etc.) cujos conteúdos são elementos descritivos do que se passa no espaço real A_R , sem, entretanto, encontrar nele referência concreta. De modo que a referência que A_R representa para qualquer B é real, porém não é concreta, em sentido metodológico. É real, porque tudo que se encontra nos diferentes conjuntos B são verdades para A_R como um ente total; todavia, A_R é abstração para qualquer B , porque os elementos descritivos que estes contêm prescindem, ou abstraem, os elementos do primeiro. Abstraindo os elementos de A_R , tornam-se impossíveis operações determinísticas que associem diretamente qualquer B com A_α , A_s , A_{ASC} .

Entretanto, dado que qualquer B , se referindo a A_R , se refere de algum modo aos seus elementos, associações probabilísticas poderão ser feitas. Com efeito, distribuições de probabilidades da ocorrência da riqueza de eventos estatísticos em associação com a riqueza de eventos espaciais digitalizados poderão ser obtidas, como se propõe em seguida. Considerem-se, inicialmente, os seguintes conjuntos contendo um painel B de informações estatísticas sobre A_R .

$$B_m = \{m \mid m = \text{usos de } a_R\} \quad (16)$$

$$B_{em} = \{e_m \mid e_m = \text{extensão do uso de } a_R \text{ associado a } m \in B_m; a_s = \sum e_m\} \quad (17)$$

$$B_{rm} = \{r_m \mid r_m = \text{expectativa de refletância do uso de } a_R \text{ associado a } m \in B_m\} \quad (18)$$

7.1. ESTATÍSTICAS E PIXELS

Considere-se, ademais, a organização dessas informações estatísticas considerando critérios igualmente válidos para as informações do *mapa de pixels*, como segue:

$$B_{ms} = \left\{ m_s \mid m_s \in B_m \wedge \left[\frac{s-1}{s}, \max(r_{ij}) \right] < r_m < \left[\frac{s}{s}, \max(r_{ij}) \right], r_m \in B_m, r_{ij} \in A_{\alpha r}, \forall s = 1, \dots, S \right\} \quad (19)$$

$$B_{msa} = \left\{ e_{ms} \mid e_{ms} \in B_m \wedge \left[\frac{s-1}{s}, \max(r_{ij}) \right] < r_m < \left[\frac{s}{s}, \max(r_{ij}) \right], r_m \in B_m, r_{ij} \in A_{\alpha r}, \forall s = 1, \dots, S, a_{ms} = \sum e_{ms} \right\} \quad (20)$$

$$B_{msp} = \left\{ p_{ms} \mid p_{ms} = \frac{e_{ms}}{\sum e_{ms}}; \sum p_{ms} = 1 \right\} \quad (21)$$

Como resultado, tem-se, a partir de pesquisa direta (censos agropecuário, por exemplo), os usos da área total a_R (16) do trato espacial A_R . Por pesquisa direta, também, sabe-se a extensão desses usos, abrigados em (17). Em (18), a partir de conhecimento laboratorial ou tácito, pesquisa amostral ou universal, direta ou indireta, têm-se valores médios presumidos de refletância para cada uso. Tais valores são condicionados por variáveis de tempo (hora do dia, dia do ano, etc.) e espaço (inclinação de incidência da luz, umidade do ar, etc.). Os valores r_m permitem a observação de uma tangência entre os sistemas de informações A (relativo aos pixels) e B (relativo às estatísticas), tanto mais nítida, quanto maior o grau de acerto da presunção sobre a refletância para m , resultado, por suposto, do nível e extensão do conhecimento associado a A_R .

Com efeito, o conjunto B_{ms} (19) contém os usos m por classes s utilizando os mesmos critérios de classificação dos pixels de A_α na organização do conjunto A_S . Assim, para a mesma classe s , se tem, ao mesmo tempo, conjuntos de informações estatísticas e conjuntos de informações relativas aos pixels. As informações estatísticas inerentemente espaciais (área plantada) informam em B_{mse} i) quanto de área (e_{ms}) requereu o modo m de utilização do espaço real, do qual se presume derivar forma física que, consideradas as experiências analisadas e os experimentos efetuados, produz padrão de refletância s nas condições médias reinantes em A_R e ii) quanto de área total ($\sum e_{ms}$) foi utilizado por todas as formas m que se enquadraram nas especificidades do padrão de refletância s (20). Disso resulta, para cada

classe s , uma distribuição de probabilidades ($p_{ms}=e_{ms}/\sum e_{ms}$) de ocorrência de m , elementos do conjunto B_{msp} (21).

7.2. ESTATÍSTICAS E SIGNIFICAÇÃO DE PIXELS

As áreas a_s , associadas às classes (genéricas) s em \dot{A}_s , submetidas às probabilidades p_m de seus usos (específicos) em B_{msp} , resultará nos elementos de um novo conjunto, como segue:

$$\dot{A}_{sm} = \{a_{sm} \mid a_{sm} = p_{sm} \cdot a_s, p_{sm} \in B_{msp}; a_s \in \dot{A}_s\} \quad (22)$$

As áreas a_{sm} , a sua vez, são a soma dos pixels a_{ij} , que pertencem a um novo conjunto A_{sm} , subconjunto do conjunto A_s . Não se sabe, ainda, que pixels exatamente são esses, mas é possível, já agora, saber em que número eles ocorrem, uma vez que

$$n_{A_{sm}} = \frac{a_{sm}}{l^2} \quad (23).$$

Sabendo o número de pixels de cada conjunto A_{sm} e que a união de todos os conjuntos A_{sm} resulta no conjunto A_s , é possível fazer o reconhecimento dos elementos de A_{sm} em A_s : isto é, é possível proceder recodificações de A_s imputando *nos seus elementos* os significados de m . Com efeito, como no capítulo 3, o conjunto de pixels de uma classe s se agrupa num conjunto A_s , resultado da união dos conjuntos A_{sm} , nele contidos, com seus complementos, conforme estabelecido em (24). No momento em que não há nenhum A_{sm} , A_{sm} é um conjunto vazio e o complemento de A_{sm} é A_s . Se se toma o elemento de menores coordenadas no complemento de A_{sm} (conf. (25)) e se verificam os elementos em torno dele em cadeia ininterrupta de vizinhança até que se contem $n_{A_{sm}}$ pixels, (conf. (26)), dispõe-se dos elementos que formam um primeiro A_{sm} . O complemento de A_{sm} , agora, de onde se elegerá o elemento de menores coordenadas, conterà apenas os elementos restantes de A_s . Assim, se formarão os sucessivos subconjuntos A_{sm} de A_s (26) até que o complemento de A_{sc} seja um conjunto vazio.

$$A_s = A_{sm} \cup \overline{A_{sm}} \quad (24)$$

$$A_T = \{a_{kp} \mid a_{kp} = a_{ij} \in \overline{A_{sm}} \wedge [k = \min(i); p = \min(j); n_{A_T} = 1]\} \quad (25)$$

$$A_{sm} = A_T \cup_{\substack{i=k+t \ (t=1, \dots, n_{A_s}) \\ j=p+t \ (t=1, \dots, n_{A_s})}}^{n_{A_{sm}}-1} A_t = \{a_{ij} \mid a_{ij} \in A_s\} \quad (26)$$

$$A_s = \cup_{m=1}^M A_{sm} \leftrightarrow \hat{A}_{sm} = \{a_{sm} \mid a_{sm} = n_{A_{sm}} \cdot l^2; \forall m = 1, \dots, M\} \quad (27)$$

Ao balanço do espaço total de AR, além das formas de agregação apresentadas, acresce as agregações por tipo de uso m nas diferentes classes s . De modo que:

$$\hat{A}_R = \{a_R \mid a_R = \sum_{i,j=1}^{I,J} a_{ij} = \sum_{s=1}^S a_s = \sum_{s,c=1}^{S,C} a_{sc} = \sum_{s,m=1}^{S,M} a_{sm} = n_{A_R} \cdot l^2; n_{\hat{A}_R} = 1\} \quad (28)$$

Acresce que se poderão desenvolver as operações de (24) a (28) tratando com os objetos discutidos em 3.2.

8. NOTAS FINAIS: MAPAS DE REALIDADE, VERDADE E INFORMAÇÃO

Acima, nos ocupamos de métodos para transformar *mapas de pixel* em *mapas de realidade* pela transferência de significados informados por *estatísticas* para os pixels de um mapa. À guisa de conclusão, as seguintes considerações:

- i. Os processos apresentados consideraram a área de referência real, do conjunto A_R , a mesma do mapa de pixels e das estatísticas. Este é o caso, por exemplo, em que se tenha pesquisa primária de um estabelecimento, ou em que se tenham as estatísticas de um município, sem georeferência das atividades. Se podem, num caso e no outro, tratar significados estruturais diretamente pela qualificação de A_R e do objeto das estatísticas.
- ii. É possível, no entanto, aplicar algoritmos derivados desse roteiro para o caso em que o *mapa de pixels* se refere à área de um estabelecimento no interior de um município e este é que é o objeto das estatísticas. Será assunto de pesquisas

futuras os problemas associados a esse uso das estatísticas do município para os estabelecimentos. O que se estabelece já aqui é que os problemas dessa ordem reduzirão com o estabelecimento de classes nas estatísticas do município que permitam enquadramento estrutural do estabelecimento. Uma elaboração das estatísticas por estrato de área dos estabelecimentos do município permitirá distribuições de probabilidades de m por faixa de refletância, como exercitamos, e por faixa de tamanho da propriedade, reduzindo as margens de erro e dando à informação significados adicionais derivados de características estruturais do objeto.

- iii. Um mapa de realidade deve ser avaliado de dois modos: pela carga de informações que carrega, esta medida pela sua *entropia relativa da informação* (h), e pelo grau de veracidade dessas informações, este medido pelo *índice kappa* de Cohen (k). A rigor, uma medida comum que pondere uma grandeza pela outra será necessária.

REFERÊNCIAS

- ALMEIDA, A. W. B. *Carajás: a guerra dos mapas*. [Belém]: Falangola, 1994. 329 p.
- CÂMARA, G., MONTEIRO, A. M. V., MEDEIROS, J. S. Representações Computacionais do Espaço: fundamentos epistemológicos da ciência da geoinformação. *Revista Geografia*, [S.l.]:UNESP, v. 28, n. 1, p. 83-96, jan./abr. 2003.
- COHEN, J. A coefficient of agreement for nominal scales. *Educational and Psychological*, [S.l.:s.n.], 1960.
- COSTA, F. A. *Formação agropecuária na Amazônia: os desafios do desenvolvimento sustentável*. Belém: NAEA, 2001. p. 299.
- _____. Capoeiras, inovações e tecnologias rurais concorrentes na Amazônia. In: COSTA, Francisco de Assis; KAHWAGE, Claudia; HURTIENNE, Thomas (Org.). *Inovação e difusão tecnológica para sustentabilidade da agricultura familiar na Amazônia: resultados e implicações do projeto SHIFT socioeconômica*. Belém: UFPA/NAEA, 2006. p. 21-58.
- _____. Dinâmica agrária e balanço de carbono na Amazônia. *Revista Economia*, Brasília, DF, v. 10, n. 1, p. 117-151, jan./abr. 2009a.
- _____. Trajetórias tecnológicas como objeto de política de conhecimento para a Amazônia: uma metodologia de delineamento. *Revista Brasileira de Inovação*, Rio de Janeiro, v. 8, n. 1, p. 35-86, jan./jun., 2009b.
- _____. Mercado e produção de terras na Amazônia: avaliação referida a trajetórias tecnológicas. *Boletim do Museu Paraense Emílio Goeldi*, Belém, v. 5, n. 1, p. 25-57, jan./abr., 2010.
- _____. *Elementos para uma economia política da Amazônia: historicidade, territorialidade, diversidade, sustentabilidade*. Belém: NAEA, 2012. 468 p. (Coleção Economia política da Amazônia. Série II – Fundamentos teórico-metodológicos; v. 2).
- _____. Heterogeneidade Estrutural, Tecnologias Concorrentes e Desenvolvimento Sustentável: uma proposição teórica para o tratamento da dinâmica agrária referida a território, com menção particular à Amazônia. In: *Boletim Regional, Urbano e Ambiental*, IPEA, Brasília-DF, v. 8, n. 2, p. 11-26, dez. 2013.
- DIAS, T. L. et al. Problemas de escala e a relação área/indivíduo em análise espacial de dados censitários. *Informática Pública*, v.4, n.1, p.89-104, 2002.
- EPSTEIN, I. *Teoria da informação*. São Paulo, Atica, 1986.

FENZL, N., HOFKIRCHNER, W. *Emergence and interaction of natural systems: the role of information, energy and matter in the perspective of a Unified Theory of Information*. Belém;Viena, NUMA;UFPA; IDTA-TU.

HABERMAS, J. *Theorie des kommunikativen handelns*. Handlungsrationality und gesellschaftliche rationalisierung. Frankfurt: Suhrkamp, 1995. p. 533.

MCCRACKEN, S. D. et al. Remote Sensing and GIS at Farm Property Level: Demography and Deforestation in the Brazilian Amazon. In: *Photogrammetric Engineering & Remote Sensing*, v. 65, n. 11, p. 1311-1320, nov. 1999.

MORAN, E. F., BRONDIZIO, E. Integrating Amazonian vegetation, land-use, and satellite data. *Bioscience*, v. 44, n. 5, p. 329-349, 1994.

MORAN, E. F. et al. Restoration of vegetation cover in the eastern Amazon. *Ecological Economics* n.18, p. 41-54 , 1996.

PONTE, M. X., DYNE, D. L. V. Sistemas Agroindustriais Integrados: Uma análise por meio da entropia da informação. *Novos Cadernos do NAEA*, v. 3, n.1, p. 47-61, jun. 2000.

SHANNON, C. E. The Mathematical Theory of Communication. *The Bell System Journal*, v. 27, p. 379-423, 623-656, 1948.

SHANNON, C.,WEAVER, W. *The Mathematical Theory of Communication*. Chicago, Illinois, University Press, 1963.

SOWA, J. A Dynamic Theory of Ontology. In: BENNET, B., C. FELLBAU (Ed.) *Formal Ontology in Information System*. Amsterdam, IOS Press.

THEIL, H. *Economics and Information Entropy*. Amsterdam, North-Holland.