

Avaliação do instrumento SAM para a etiquetagem colaborativa em ambiente web de imagens segundo critérios afetivos

Wellton C. Costa
Universidade Tecnológica
Federal do Paraná - UTFPR
CPGEI/LABIEM
welltonutfpr@gmail.com

Álvaro R. Cantieri
Universidade Tecnológica
Federal do Paraná - UTFPR
CPGEI/LABIEM
alvaro.cantieri@yahoo.com.br

Gustavo B. Borba
Universidade Tecnológica
Federal do Paraná - UTFPR
CPGEI/LABIEM
gustavobborba@utfpr.edu.br

Humberto R. Gamba
Universidade Tecnológica
Federal do Paraná - UTFPR
CPGEI/LABIEM
humberto@utfpr.edu.br

ABSTRACT

In a visual information retrieval process that considers the images semantics, several levels of query can emerge, from simple statements as “find pictures with a boat”, to more abstract ones, such as “find pictures depicting happy atmospheres”. Usually, the abstract level refers to the affective or emotional content of the images and is considered a relevant dimension in which users specify their queries. Clearly, due to its inherent complexity, affect is difficult to model and consequently quite difficult to be handled by a retrieval system. In this work, we present a preliminary evaluation of the Self Assessment Manikin (SAM) method for the affective description of images, in a collaborative tagging environment. Images from the International Affective Picture System (IAPS) dataset were used in the study. The results show that the affective dimensions valence and arousal and the SAM method can be efficiently used for tagging and further retrieval of images.

Categories and Subject Descriptors

H.3.3 [Information Storage and Retrieval]: Information Search and Retrieval

General Terms

Experimentation, Human Factors

Keywords

Affective tagging, image retrieval, valence/arousal, SAM, IAPS

RESUMO

Em uma recuperação de imagens a partir de critérios semânticos a especificação da busca pode ocorrer em diferentes níveis. A especificação “encontre imagens contendo um barco”, por exemplo, pode ser considerada simples e objetiva se comparada com “encontre imagens ilustrando um clima alegre”, nitidamente mais subjetiva e abstrata. Em geral, este nível abstrato refere-se ao conteúdo afetivo ou emocional das imagens e mostra-se bastante relevante no contexto da recuperação semântica de imagens. No entanto, devido à sua complexidade, o nível afetivo é difícil de ser modelado e, conseqüentemente, difícil de ser tratado em um sistema de recuperação de imagens. Neste trabalho, apresenta-se uma avaliação preliminar do instrumento *Self Assessment Manikin* (SAM) para a etiquetagem colaborativa em ambiente web de imagens segundo critérios afetivos. Os resultados destes testes com voluntários e imagens da biblioteca *International Affective Picture System* (IAPS) demonstraram que as dimensões *valência* e *alerta*, em conjunto com o instrumento SAM, podem ser explorados como um recurso eficiente para a etiquetagem e futura recuperação de imagens, considerando o critério afetivo das mesmas.

1. INTRODUÇÃO

As recentes facilidades para a geração e compartilhamento de conteúdo visual têm contribuído para o grande crescimento deste tipo de material na web. Com isto, o estudo de métodos eficientes para a recuperação destes dados mostra-se cada vez mais relevante. Uma estratégia bastante difundida para a *recuperação de conteúdo visual* (RCV) é a *recuperação de imagens com base no conteúdo* (CBIR - *content-based image retrieval*). No CBIR, as imagens são indexadas a partir das suas características de baixo nível, como cor, forma e textura [1].

Sistemas exclusivamente CBIR, porém, não levam em consideração um importante aspecto da RCV: o significado ou semântica das imagens. Os usuários geralmente estão interessados no significados das imagens, e não em seus atributos de baixo nível, relacionados aos pixels (*raw data*). A discrepância entre aquilo que os usuários têm em mente ao

executar uma busca e aquilo que o sistema de RCV é capaz de extrair das imagens é denominado de *gap semântico* [2, 3].

Apesar dos constantes e significativos avanços, uma solução geral para o gap semântico ainda não é uma realidade [1, 2]. Além disso, inúmeros autores importantes são cuidadosos ao referirem-se ao gap semântico, sugerindo que uma solução geral dificilmente será obtida, devido a fatores como a subjetividade naturalmente presente na interpretação de uma imagem [4].

Neste contexto, a comunidade de RCV tem gradualmente direcionado esforços para o desenvolvimento de soluções práticas, ao invés de exclusivamente para o desenvolvimento de uma solução definitiva para o problema do gap semântico. As abordagens para isto envolvem, por exemplo, a construção de sistemas dedicados a aspectos específicos das imagens [5] e o uso de metadados gerados pelos usuários, em ambientes de etiquetagem colaborativa [4, 6].

Considerando-se a semântica das imagens, podem ocorrer diferentes tipos de especificação para a busca, como por exemplo, especificações simples como “encontre imagens contendo um barco”, até mais abstratas, como “encontre imagens contendo um clima positivo”. Geralmente, uma especificação abstrata refere-se ao conteúdo *afetivo* ou *emocional* das imagens, e trata-se de um critério importante para os usuários [7]. Uma das maneiras de realizar a medição do aspecto afetivo de uma imagem é chamada de *representação dimensional da emoção* [8]. Um instrumento utilizado para este fim é uma ferramenta denominada SAM [9].

Este artigo tem como objetivo mostrar os resultados preliminares de um estudo que usa uma ferramenta web colaborativa, chamada *Get your mood*, para etiquetar imagens segundo o critério afetivo através do SAM. O presente artigo também faz um estudo comparativo entre os dados obtidos pela ferramenta web proposta com os valores dimensionais das emoções disponibilizados pela biblioteca IAPS [10], de forma a verificar a viabilidade da etiquetagem realizada através da web.

2. REPRESENTAÇÃO DIMENSIONAL DA EMOÇÃO

Na representação dimensional da emoção, utilizam-se duas dimensões independentes: uma chamada de *valência* (variando de prazeroso para não-prazeroso), e outra chamada de *alerta*¹ (variando de excitado para relaxado). Estas duas dimensões podem ser combinadas em um gráfico que demonstra a reação afetiva do usuário para cada imagem avaliada. No procedimento descrito por [10], voluntários em ambientes de laboratório avaliam um conjunto de imagens utilizando regras e procedimentos previamente definidos, buscando identificar suas reações afetivas para cada uma dessas imagens.

2.1 Instrumento SAM

Um instrumento bastante utilizado para a obtenção dos valores de valência e alerta é o SAM [9]. Este instrumento é

¹Tradução usualmente utilizada na literatura brasileira para o termo *arousal* da língua inglesa [11].

formado por três grupos de bonecos impressos em papel, representando a resposta do usuário em cada dimensão da emoção. O primeiro grupo de bonecos representa a dimensão *valência*, o segundo grupo representa a dimensão *alerta* e o terceiro grupo representa a dimensão *dominância*. Usualmente, os estudos realizados com essa ferramenta só utilizam as duas primeiras dimensões. O voluntário procede no teste marcando um X com um lápis sobre o boneco de cada grupo que melhor descreve sua reação. O SAM possui boas qualidades psicométricas e é um instrumento confiável e válido porque as classificações de valência e alerta estão altamente correlacionadas com as medidas de avaliação afetiva obtidas na Escala Semântica Diferencial, elaborada por Mehrabian & Russel [12, 13]. A Figura 1 apresenta os bonecos SAM para as dimensões valência e alerta.

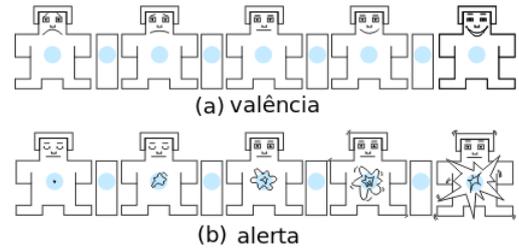


Figura 1: Instrumento SAM.

2.2 Biblioteca de imagens IAPS

A IAPS é uma biblioteca de imagens com centenas de fotografias coloridas que representam vários aspectos da vida real (esportes, moda, paisagens, violência, etc.), com o objetivo de evocar diferentes estados emocionais. Esta biblioteca é disponibilizada para estudos com pacientes e, segundo suas regras de uso, não podem ser divulgadas, exceto durante os experimentos. As avaliações realizadas nas imagens da IAPS utilizam o SAM. Assim, além dos arquivos de imagens, a biblioteca IAPS disponibiliza também os valores de média e desvio padrão das dimensões emocionais associadas a cada imagem [14]. A Figura 2 mostra dois exemplos de imagens dispostas no gráfico valência *versus* alerta. Estas imagens não são provenientes da IAPS.

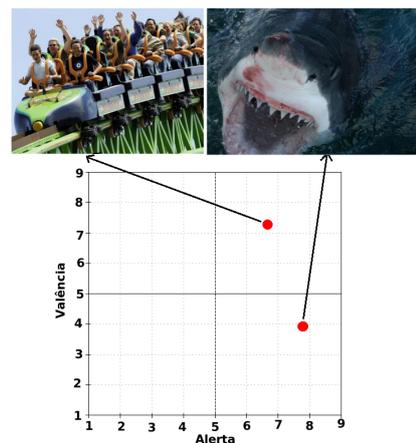


Figura 2: Exemplo de duas imagens no gráfico valência *versus* alerta.

3. METODOLOGIA

A metodologia utilizada neste trabalho consiste na criação de uma ferramenta web para a aplicação dos testes de resposta afetiva em um conjunto de imagens e sua comparação estatística com os valores disponibilizados pela biblioteca IAPS. São descritos a seguir a ferramenta desenvolvida e também os testes estatísticos realizados.

3.1 Ferramenta Get your mood

Para possibilitar a aplicação do SAM na internet, foi necessária a criação de uma ferramenta web. Esta ferramenta, chamada *Get your mood*, foi implementada na linguagem PHP, para o sistema de login; MySQL, para o armazenamento dos dados; e ajax, para o armazenamento de resultados assíncronos. O servidor web utilizado foi o Apache e o sistema operacional utilizado foi o Debian GNU/Linux 5.0 (Lenny).

Ao se conectar ao site www.getyourmood.com, o usuário recebe uma mensagem inicial e dois campos solicitando sua identificação e senha. Tanto a identificação quanto a senha são fornecidos através de métodos *offline* para os voluntários previamente cadastrados. Isto é importante para garantir que as imagens da IAPS não sejam levemente disponibilizadas na internet, já que as mesmas são protegidas e disponibilizadas apenas para estudos. Como mencionado anteriormente, as imagens da IAPS não podem ser divulgadas, para não serem previamente conhecidas por eventuais voluntários de experimentos científicos. Por este motivo, a Tabela 1 apresenta apenas a descrição textual das imagens utilizadas. Neste estudo, foram utilizadas dez imagens desta biblioteca.

Tabela 1: Descrição das imagens do experimento

n°	Descrição
1	Filhotes de cobra
2	Barata
3	Três filhotes de cão
4	Homem de chapéu, barba e óculos tocando violão
5	Mulher idosa sorrindo
6	Mulher saltando de bungee jump sobre uma ponte
7	Crianças em um tobogã de água
8	Barras de ouro empilhadas
9	Crânios humanos empilhados
10	Pato preso no óleo em um lago

Antes de iniciar o experimento, o usuário recebe uma rápida explicação sobre o funcionamento do mesmo e, em seguida, passa para a visualização da primeira imagem. Esta imagem é mostrada durante seis segundos para o usuário e, na sequência, o instrumento SAM é apresentado. O usuário clica nos níveis de valência e alerta que melhor descrevem a imagem segundo a sua interpretação. Ao fazer isso, o sistema automaticamente envia o resultado para o banco de dados e outra imagem é mostrada. O processo se repete até que o usuário complete a avaliação de cinco imagens, o que constitui uma sessão completa. A Figura 3 apresenta duas telas da *Get your mood* ilustrando este processo.

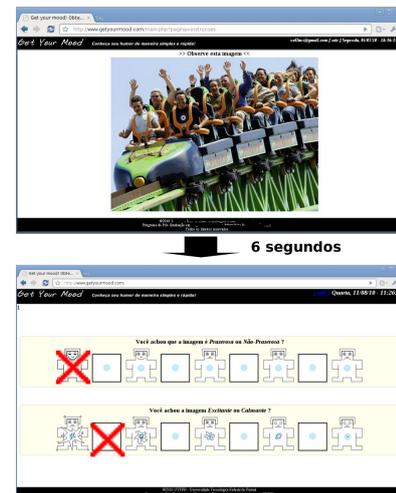


Figura 3: Duas telas da ferramenta *Get your mood*. Primeiramente, a imagem sendo mostrada ao usuário. Após seis segundos são apresentados os bonecos SAM para a etiquetagem.

3.2 Aplicação dos testes e tabulação dos resultados

Para a aplicação dos testes, uma média de 30 pessoas participou, respondendo sobre sua sensação para cada uma das dez imagens presentes na ferramenta. Com as respostas dos voluntários tabuladas para cada imagem, foram calculados os valores de média e desvio padrão da valência e do alerta para cada uma delas. Estes valores são mostrados nas Tabelas 2 e 3. Também foi realizado um teste de comparação estatística entre os resultados obtidos através do *Get your mood* e os resultados disponibilizados pela IAPS. Este teste consiste na verificação da existência, para cada uma das médias de valência e alerta das imagens utilizadas, de diferenças estatísticas significativas entre estas e as médias da IAPS. Para isso, foi utilizado um teste de hipóteses de comparação de médias com variância populacional conhecida, para distribuições normais. A aproximação da distribuição normal para os valores de valência e alerta respondidos pelos voluntários é considerada válida neste caso. A variância da população corresponde à variância definida pela IAPS. A Equação 1 [15] descreve o teste.

$$Z_0 = \frac{\bar{X} - \bar{x}}{\frac{S^2}{\sqrt{n}}}, \quad (1)$$

onde Z_0 é a variável padronizada do teste estatístico, \bar{X} é a média dos dados do IAPS, \bar{x} é a média dos dados do *Get your mood*, S^2 é a variância dos dados do IAPS e n é o número de avaliações realizadas por imagem.

Considerando-se um nível de significância de 5%, rejeita-se a hipótese nula (média das imagens avaliadas através do *Get your mood* igual às médias das imagens da IAPS) caso a variável padronizada Z_0 esteja fora da faixa $[-1,94 \dots +1,94]$.

4. RESULTADOS

Os resultados obtidos, apresentados nas Tabelas 2 e 3, mostram que, para todas as imagens avaliadas, a hipótese nula foi confirmada para o nível de significância definido no teste. Ou seja, a avaliação estatística das médias das dimensões valência e alerta, para cada imagem, demonstraram que não existem diferenças significativas entre as médias provenientes da IAPS e as médias obtidas através do *Get your mood*. Ainda, as Tabelas 2 e 3 mostram que os valores de desvio padrão das imagens testadas ficaram bastante próximos dos valores de desvio padrão disponibilizados pela IAPS. Como estes resultados são preliminares, um número maior de avaliações, sobre um conjunto maior de imagens da IAPS, deve ser realizado. Estas avaliações estão em andamento no momento.

Tabela 2: Comparação entre os resultados obtidos através do *Get your mood* e os valores disponibilizados pela IAPS, para a dimensão valência

n°	<i>Get your mood</i>		IAPS		Z ₀
	Média	(DP)	Média	(DP)	
1	3,25	(1,64)	3,84	(1,61)	-1,22
2	3,68	(1,85)	4,00	(1,47)	-0,46
3	8,34	(1,12)	8,10	(1,40)	1,047
4	5,20	(1,80)	4,89	(1,78)	0,49
5	6,91	(1,91)	6,73	(1,20)	0,23
6	6,48	(2,18)	7,32	(1,84)	-0,82
7	7,58	(1,63)	7,64	(1,66)	-0,11
8	6,96	(1,64)	7,12	(1,96)	-0,29
9	3,67	(1,86)	3,25	(2,51)	0,54
10	2,21	(1,93)	3,00	(2,11)	-1,10

Tabela 3: Comparação entre os resultados obtidos através do *Get your mood* e os valores disponibilizados pela IAPS, para a dimensão alerta

n°	<i>Get your mood</i>		IAPS		Z ₀
	Média	(DP)	Média	(DP)	
1	4,58	(2,38)	5,20	(2,25)	0,68
2	4,60	(2,10)	4,77	(2,44)	0,14
3	4,87	(1,74)	5,41	(1,98)	0,75
4	4,00	(1,64)	4,05	(1,92)	0,07
5	4,00	(1,60)	4,00	(2,10)	0,00
6	7,23	(1,97)	6,99	(2,35)	-0,20
7	7,14	(2,10)	5,79	(2,26)	-1,39
8	5,64	(1,89)	5,60	(2,40)	-0,03
9	5,50	(2,67)	4,55	(2,02)	-1,04
10	5,41	(2,14)	5,50	(2,52)	0,073

5. CONCLUSÃO

Os resultados obtidos até o momento apontam para a viabilidade da etiquetagem de imagens segundo critérios afetivos, em ambientes web colaborativos, utilizando as dimensões valência e alerta juntamente com o instrumento SAM. Mais avaliações, com o número maior de usuários e de imagens, estão em andamento. Um dos pontos a se destacar na avaliação afetiva em ambientes colaborativos via web é a sua penetração em grandes grupos de voluntários de maneira simples e direta. Os resultados apresentados nesta etapa preliminar do trabalho não têm o objetivo de diminuir o gap semântico diretamente, mas de avaliar a possibilidade de criação de um

método para a etiquetagem de imagens segundo critérios afetivos. Uma vez comprovada a eficiência deste método, um grande número de imagens pode ser etiquetado colaborativamente, permitindo a construção de uma ferramenta de recuperação de imagens a partir destas etiquetas, o que efetivamente irá contribuir para a diminuição do gap semântico.

6. REFERÊNCIAS

- [1] R. Datta, D. Joshi, J. Li e J. Wang. Image retrieval: Ideas, influences, and trends of the new age. *ACM Comput. Surv.*, 40(2):1-60, 2008.
- [2] A. Jaimes. Human factors in automatic image retrieval system design and evaluation. *IS&T SPIE Proceedings of SPIE - Internet Imaging VII*, volume 6061, 2006.
- [3] Y. Liu, D. Zhang, G. Lu e W. Ma. A survey of content-based image retrieval with high-level semantics. *Pattern Recogn.* 40(1):262-282, 2007.
- [4] C. Jorgensen. Image access, the semantic gap, and social tagging as a paradigm shift. *Proceedings 18th Workshop of the American Society for Information Science and Technology Special Interest Group in Classification Research*, 2007.
- [5] Like.com Visual Search. <http://www.like.com/>. Acessado em julho de 2010.
- [6] C. A. Basso e S. R. Silva. Uma proposta para a evolução de ontologias a partir de folksonomias. *Companion Proceedings of the Brazilian Symposium on Multimedia and the Web, WebMedia '08 ACM*, p.197-200, Vila Velha, Espírito Santo, Brazil, 2008.
- [7] J. Eakins. Towards intelligent image retrieval. *Pattern Recogn.*, 35(1):3-14, 2002.
- [8] M. Bradley e P. Lang. Measuring emotion: behavior, feeling and physiology. *R. Lane, L. Nadel. Cognitive Neuroscience of Emotion*, Oxford University, New York, 2000.
- [9] P. Lang. Behavioral treatment and biobehavioral assessment: computer applications. *J.B. Sidowski, J.H. Johnson, T.A. Williams, Technology in mental health care delivery systems.*, Norwood, JR: Ablex. p.119-37, 1980.
- [10] P. Lang, M. Bradley e B. Cuthbert. International affective picture system (IAPS): Affective ratings of pictures and instruction manual. *Technical Report A-6, University of Florida*, Gainesville, 2005.
- [11] R. Ribeiro, S. Pompéia e O. Bueno. Normas brasileiras para o International Affective Picture System (IAPS): comunicação breve. *Revista de Psiquiatria do Rio Grandedo Sul*, 26:190-194, 2004.
- [12] A. Mehrabian, A. e J. Russell. An approach to environmental psychology. MIT Press, Cambridge, 1974.
- [13] M. Bradley e P. Lang. Cries, screams, and shouts of joy: Affective responses to environmental sounds [Abstract]. *Psychophysiology*, 31(1):29, 1994.
- [14] P. Lang, M. Bradley e B. Cuthbert. International Affective Picture System (IAPS): Instruction manual and affective ratings. *Technical Report A-4, University of Florida*, Gainesville, 1999.
- [15] D. Montgomery e G. Runge. Estatística Aplicada à Engenharia. Livros Técnicos e Científicos, Rio de Janeiro, 2001.