
Estudo da Percepção Humana e Mineração de Preferências Contextuais na Recomendação de Imagens para o Usuário

Claudianne Myllene Morais de Almeida



UNIVERSIDADE FEDERAL DE UBERLÂNDIA
FACULDADE DE COMPUTAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

Uberlândia
2015

Claudianne Myllene Morais de Almeida

**Estudo da Percepção Humana e Mineração de
Preferências Contextuais na Recomendação de
Imagens para o Usuário**

Dissertação de mestrado apresentada ao Programa de Pós-graduação da Faculdade de Computação da Universidade Federal de Uberlândia como parte dos requisitos para a obtenção do título de Mestre em Ciência da Computação.

Área de concentração: Banco de Dados e Imagens

Orientador: André Ricardo Backes

Uberlândia
2015

Dados Internacionais de Catalogação na Publicação (CIP)
Sistema de Bibliotecas da UFU, MG, Brasil.

A447e
2016 Almeida, Claudianne Myllene Moraes de, 1992-
Estudo da percepção humana e mineração de preferências
contextuais na recomendação de imagens para o usuário / Claudianne
Myllene Moraes de Almeida. - 2016.
78 f. : il.

Orientador: André Ricardo Backes.
Dissertação (mestrado) - Universidade Federal de Uberlândia,
Programa de Pós-Graduação em Ciência da Computação.
Inclui bibliografia.

1. Computação - Teses. 2. Percepção visual - Teses. I. Backes,
André Ricardo. II. Universidade Federal de Uberlândia, Programa de
Pós-Graduação em Ciência da Computação. III. Título.

CDU: 681.3

Dedico este trabalho a minha mãe, Maria Dalva, minha maior incentivadora que sempre me ajudou e acreditou que seria capaz de realizar mais um sonho meu.

Agradecimentos

Agradeço primeiramente a Deus, que sempre me deu força para eu nunca desistir. E além disso, ele colocou do meu lado pessoas muito especiais. Algumas delas me encorajaram e me ajudaram a chegar até onde cheguei.

Agradeço a minha família que sempre me deu todo o carinho e apoio necessário para que esse meu sonho fosse realizado. Em especial a minha mãe, Maria Dalva, e aos meus irmãos, Michelle, Mariane e Herus pelo amor incondicional.

Agradeço a meus amigos Bruno, Emília, Walter, Maria Cristina, Joicy, Sara, Pedro, Maicon, Edinilson, Regis, Marcos, Guilherme e Juan por todo carinho, companheirismo, paciência e compreensão que sempre tiveram comigo. As minhas amigas de sempre, Tati e Josiely, por me valorizarem tanto como pessoa e sempre querer o meu bem. A todos vocês obrigada pela amizade.

Agradeço de forma especial as minhas amigas Franciny, Fabíola e Fernanda, que estiveram ao meu lado durante os momentos cruciais do mestrado e da elaboração dessa dissertação. Vocês foram simplesmente essenciais.

A todos os professores que contribuíram para a minha formação, que me apoiaram e me estimularam a chegar até aqui.

Agradeço, também, a CAPES pelo apoio financeiro nesse período do mestrado.

Agradeço a professora Sandra de Amo por sua dedicação, atenção, competência e sugestões durante o mestrado. Principalmente, por ter me mostrado os primeiros passos da pesquisa científica.

Ao meu orientador, André Backes, pela atenção nas revisões, sugestões e confiança, fatores fundamentais para a conclusão deste trabalho.

Aos meus amigos de laboratório do LSI pelo apoio nos momentos difíceis e por todo o aprendizado que eu conquistei ao lado de vocês.

A todos os meus amigos que conheci antes e durante o mestrado.

Ter amigos e professores que pensam de forma tão diferente foi enriquecedor para minha formação. Agradeço a todos que me rodeia que, apesar de me desorientar às vezes, me ajuda a enxergar diferentes olhares sobre a mesma realidade. Não tenho palavras para

agradecer tudo o que vocês fizeram por mim. Obrigada de coração.

“Que os vossos esforços desafiem as impossibilidades, lembrai-vos de que as grandes coisas do homem foram conquistadas do que parecia impossível.”
(Charles Chaplin)

Resumo

Este trabalho apresenta um estudo sobre o uso da informação da percepção visual na recomendação de imagens. Assim, foi proposto um método que faz o agrupamento dos usuários através de sua percepção visual e de sua similaridade. Esse método foi denominado de *VP-Similarity*.

O *VP-Similarity* foi implementado por meio da extensão do sistema *PrefRec* para utilização de dados de percepção visual em conjunto com dados de preferência e de itens na tarefa de recomendação, formando assim o *VP-PrefRec*. Para validação do sistema foi criada uma base de dados. Essa base de dados contém dados de preferência e dados de percepção visual dos usuários. Além disso, o *VP-Similarity* também foi aplicado sobre um sistema de recomendação social, fornecendo uma rede de percepção visual. O objetivo dessa rede é minimizar o problema do *cold start* do usuário, existente na maioria dos sistemas de recomendação.

A finalidade da proposta deste trabalho é mostrar que a recomendação de imagens com dados de itens e de preferências contextuais para um usuário alvo tem melhor qualidade quando são considerados apenas os usuários que têm percepções visuais similares a esse usuário alvo, considerando algumas avaliações prévias desse usuário alvo. E também mostrar que é possível melhorar a qualidade das recomendações de imagens para um novo usuário que nunca tenha feito uso do sistema (*cold start*), concentrando-se apenas em usuários que visualmente percebam as imagens de forma similar.

Palavras-chave: percepção visual, sistema de recomendação, rede de percepção visual, *cold start*.

Abstract

This work presents a study on the use of information of visual perception in the image recommendation. Thus, it proposed a method that makes the grouping of users through your visual perception and their similarity. This method was denominated *VP-Similarity*.

The *VP-Similarity* is implemented by extending the system *PrefRec* to use of visual perception data in conjunction with preference data item and the recommendation task, thereby forming the *VP-PrefRec*. For the validation of the system a database was created. This database contains preference and visual perception data of users. In addition, *VP-Similarity* was also applied on a social recommendation system, providing a visual perception network. The purpose of this network is to minimize the user cold start problem, which in most recommender systems.

The objective of purpose of this paper is to show that the recommendation of images with data items and contextual preferences for a target user has better quality when considering only those users who have similar visual perceptions that target user, considering some prior evaluations of that user target. And it also shows that it is possible to improve the quality of images of recommendations for a new user who has never made use of the system (cold start), focusing only on users to visually perceive the similar images.

Keywords: visual perception, recommendation system, visual perception network, cold start.

Lista de ilustrações

Figura 1 – Pintura sendo observada por dois grupos de pessoas.	20
Figura 2 – Um novo usuário entra no sistema.	20
Figura 3 – Tipos de dados de entrada do <i>VP-PrefRec</i>	21
Figura 4 – Representação de um cubo RGB. Fonte: (GONZALEZ; WOODS, 2010). .	30
Figura 5 – Exemplo de uma imagem e seu histograma. Fonte: (GONZALEZ; WOODS, 2010).	31
Figura 6 – Hierarquia das técnicas de extração de formas. Fonte: (ZHANG; LU, 2004).	31
Figura 7 – Esquema geral do <i>PrefRec</i> . Adaptado de: (AMO; OLIVEIRA, 2014). .	33
Figura 8 – Exemplo de uma <i>Mpref</i> de um usuário: (a) Avaliações de um usuário. (b) Matriz retornada pelo Módulo 1, correspondendo as preferências do usuário. Os (*) nas tabelas significam que determinado item não foi avaliado pelo usuário.	33
Figura 9 – Tipos de dados dos sistemas de recomendação.	35
Figura 10 – Visão geral da principal proposta do trabalho.	41
Figura 11 – Exemplo da clusterização por percepção visual. Em destaque de azul estão os usuários do <i>cluster</i> 1 e os de vermelho são os usuários do <i>cluster</i> 2.	43
Figura 12 – Arquitetura geral do <i>VP-PrefRec</i> . As áreas destacadas de vermelho são as contribuições deste trabalho.	44
Figura 13 – Algumas imagens da base de dados.	47
Figura 14 – Atributos das imagens e a quantidade de variações de cada um dos atributos. Essas variações vão influenciar na hora do <i>VP-PrefRec</i> construir os modelos de recomendação.	47
Figura 15 – Atributos de cada imagem.	48
Figura 16 – Mapa de calor do teste realizado na Figura 15.	49
Figura 17 – Exemplos de uma imagem dividida em q partes: (a) $q = 4$; (b) $q = 9$; (c) $q = 16$; (d) $q = 25$;	53

Figura 18 – Vetores criados de cada imagem avaliada por um usuário.	53
Figura 19 – Cálculo do Coeficiente de Silhueta para $k = 2$: (a) $q = 4$ e (b) $q = 9$.	55
Figura 20 – Cálculo do Coeficiente de Silhueta para $k = 3$: (a) $q = 4$ e (b) $q = 9$.	56
Figura 21 – Resultado do NDCG = 5.	60
Figura 22 – Resultado do NDCG = 10.	60
Figura 23 – Resultado do NDCG = 15.	61
Figura 24 – Resultado do NDCG = 20.	61
Figura 25 – Variação do <i>threshold</i> da métrica <i>similarity</i>	62
Figura 26 – Resultado do NDCG = 5.	63
Figura 27 – Resultado do NDCG = 10.	63
Figura 28 – Resultado do NDCG = 15.	64
Figura 29 – Resultado do NDCG = 20.	64
Figura 30 – Círculo cromático representando as cores frias e quentes.	65
Figura 31 – Barra de cores representadas em graus.	66
Figura 32 – Dados dos itens.	67
Figura 33 – Comparação Acurácia com o uso de atributos imagens e sem o uso desses atributos.	67

Sumário

1	INTRODUÇÃO	17
1.1	Motivação	19
1.2	Objetivos	22
1.3	Hipótese	22
1.4	Organização da Dissertação	22
2	FUNDAMENTOS TEÓRICOS	25
2.1	Sistema de Recomendação	25
2.1.1	Técnica de Filtragem Colaborativa	26
2.1.2	Técnica Baseada em Conteúdo	27
2.1.3	Abordagem Híbrida	28
2.2	Percepção visual	28
2.3	Extração de características de imagem	29
2.4	PrefRec	32
3	REVISÃO BIBLIOGRÁFICA	35
3.1	Sistemas de Recomendação	37
3.2	Sistemas de Recomendação que utilizam rastreadores de olhar	39
4	PROPOSTA	41
4.1	Método <i>VP-Similarity</i>	41
4.2	<i>VP-PrefRec</i> : uma extensão do <i>PrefRec</i>	43
4.3	Inferindo uma rede de percepção visual	45
4.4	Criação da base de dados	46
5	RESULTADOS EXPERIMENTAIS	51
5.1	Pré-Processamento da base de dados	51
5.2	Divisão das imagens em <i>q</i> -partes	51

5.3	<i>VP-Similarity</i> : algoritmo de clusterização dos usuário por percepção visual	54
5.3.1	Seleção da melhor medida de distância	54
5.3.2	Seleção da melhor configuração de quadrantes	57
5.4	Desempenho do <i>VP-PrefRec</i>	59
5.5	Influência da inferência da rede de percepção visual	61
5.6	Influência dos atributos de imagens	64
6	CONCLUSÃO	69
6.1	Principais Contribuições	70
6.2	Trabalhos Futuros	70
6.3	Contribuições em Produção Bibliográfica	71
	REFERÊNCIAS	73

CAPÍTULO 1

Introdução

Os avanços tecnológicos e a facilidade de acesso aos meios de comunicação estimularam a sociedade na busca de informações. Essas informações podem ser encontradas na *Internet* (por meio de notícias, imagens, vídeos, por exemplo), programas de televisão ou até mesmo fornecidas através das pessoas (conversas). A *Internet*, principalmente, tem permitido as pessoas a conveniência de poder conversar, buscar e trocar informações de qualquer lugar do mundo (KIESLER, 2014).

O uso de sistemas de transações comerciais por meio da *Internet* tem aumentado cada vez mais (BARNES; HUNT, 2013; HE; CHI, 2013). Com a importância das transações, os usuários possuem acesso a uma grande quantidade e diversidade de produtos que são ofertados nesse mundo virtual. Esse grande fluxo de informação e de produtos faz com que os usuários tenham dificuldade de encontrar rapidamente o que eles procuram. Com isso, é interessante estudar e propor maneiras de ajudar os usuários a encontrar conteúdos ou produtos de seu interesse (HUANG; CHUNG; CHEN, 2004; BHATT; PATEL; GAUDANI, 2014).

Para resolver o problema da busca de conteúdos, surgiram os sistemas de recomendação que têm como propósito prover informações personalizadas aos usuários através da sua capacidade de reconhecer preferências. De acordo com Shani e Gunawardana (2011), é possível encontrar sistemas de recomendação em diversas aplicações modernas, que apresentam uma enorme coleção de itens aos usuários. Esses sistemas identificam itens com base em informações de preferência pessoal, que têm uma grande chance de atender os gostos dos usuários (JANNACH et al., 2010; ADOMAVICIUS; TUZHILIN, 2011). *Google*, *Youtube*, *Netflix* e *Amazon* são exemplos de empresas que usam esses tipos de sistemas com o propósito de melhorar a satisfação dos seus clientes (LINDEN; SMITH; YORK, 2003; LIU; DOLAN; PEDERSEN, 2010; DAVIDSON et al., 2010; BERRY et al., 2010).

Dentre as técnicas de sistema de recomendação, duas são as mais utilizadas: filtragem colaborativa (do inglês *Collaborative Filtering*, CF) e recomendação baseada em conteúdo (do inglês *Content-Based*, CB). A técnica de filtragem colaborativa é utilizada para reco-

mendar itens para um usuário levando em consideração itens avaliados por outros usuários com preferência similares (SARWAR et al., 2001; YAO et al., 2013; HERLOCKER et al., 2004; ADOMAVICIUS; TUZHILIN, 2005; CANDILLIER; MEYER; BOULLÉ, 2007). Como exemplo, podemos considerar um cenário de recomendação de filmes. Para fazer a recomendação de um filme para um usuário x que prefere de filmes de suspense, o sistema fará uma busca na sua base de dados por perfis de usuários que também gostam desse tipo de filme e, assim, utilizará as opiniões desses outros usuários para recomendar filmes para x . Um dos desafios na utilização dessa técnica é o problema *cold-start de item* (SCHEIN et al., 2002; LAM et al., 2008; QIU et al., 2011; SHANI; GUNAWARDANA, 2011), em que um item é recomendado a algum usuário somente se esse item foi avaliado por outros usuários similares ao usuário ativo.

A técnica de recomendação baseada em conteúdos, por outro lado, sugere itens similares ao que o usuário avaliou (HERLOCKER, 2000; ADOMAVICIUS; TUZHILIN, 2005; WEI; HUANG; FU, 2007; LOPS; GEMMIS; SEMERARO, 2011; BARRILERO et al., 2011). Dessa maneira, com base nas características dos itens que ainda não foram avaliados é possível recomendá-los. Um exemplo simples da utilização dessa técnica é um sistema de recomendação de filmes e um usuário x . Esse usuário assistiu vários filmes de comédia e alguns desses filmes têm a participação do ator Adam Sandler. Com base nessas informações, o sistema recomendará para tal usuário outros filmes de comédias e, dentre eles, alguns contendo a participação desse ator. Contudo, os sistemas de recomendação que utilizam essa técnica são super especializados, ou seja, os itens que não são semelhantes aos que foram previamente avaliados dificilmente serão recomendados (LOPS; GEMMIS; SEMERARO, 2011). Além disso, a recomendação baseada em conteúdo sofre com o problema de *cold start do usuário* (LAM et al., 2008; LOPS; GEMMIS; SEMERARO, 2011; BERNARDI et al., 2015). Esse problema faz com que os sistemas não recomendem itens para novos usuários devido a falta de avaliações dos mesmos.

Com o propósito de lidar com as limitações das técnicas de filtragem colaborativa e baseada em conteúdo, surgiram as abordagens de recomendação híbrida. Essas abordagens consistem em combinar as duas técnicas a fim de aproveitar os benefícios de cada uma delas e, com isso, criar sistemas de recomendação que possam atender melhor as particularidades dos usuários (BALABANOVIĆ; SHOHAM, 1997; BURKE, 2005; BURKE, 2007; WOERNDL; SCHUELLER; WOJTECH, 2007; OLIVEIRA; AMO, 2013; AMO; OLIVEIRA, 2014).

Além dos sistemas de recomendação, a percepção visual das pessoas exerce grande influência na compreensão dos dados e da sua utilidade. O sistema visual é um dos principais sentidos para interação dos seres vivos com o ambiente. Conforme (TORTORA; DERRICKSON, 2012), a visão é o sentido que responde a estímulos de luz. Assim, uma particularidade interessante na visão biológica é a seleção de informações relevantes de uma cena, pois o volume de informações que estimulam o campo visual dos sistemas

biológicos é grande e, com seleção, a quantidade de informações a serem processadas é reduzida (TSOTSOS, 1990). Wan, Jin e Yue (2009) destacam que pesquisas da psicologia visual evidenciam que quando se analisa uma imagem as pessoas não têm o mesmo interesse sobre ela. Determinadas partes de uma imagens podem ser mais interessantes do que outras para diferentes pessoas. Assim, a percepção visual se torna uma grande aliada na tarefa de análise de dados e na personalização de informação para cada usuário.

Dentro desse contexto, o objetivo deste trabalho foi propor um sistema de recomendação híbrido de imagens baseado na percepção humana e nas preferências dos usuários. Atualmente existem poucos trabalhos dedicados à utilização de percepção visual de imagens no processo de recomendação. Assim, a ideia geral deste trabalho é comprovar que é possível melhorar a recomendação de imagens com o uso da percepção visual de usuários por meio de um *eye tracker*. Isso se assemelha e estende a abordagem de Umemoto et al. (2012) que encontrou como barreira a existência de imagens e animações em páginas para seu sistema de recomendação.

Para tanto, foi criado uma base de dados usando imagens de obras de arte e percepções visuais de usuários. As percepções visuais dos usuários foram obtidas por um rastreador ocular (*eye tracker*). Além disso, a extração de características das imagens foi feita com o intuito de descobrir, através das percepções visuais e das preferências dos usuários, o que atrai a atenção das pessoas fazendo com que elas gostem ou não de uma imagem. Por fim, também verificamos se a extração das características das imagens contribuía para melhorar a recomendação do sistema. Espera-se com o uso de dados sobre percepção humana aprimorar as recomendações de imagens. Com o intuito de avaliar se houveram melhoras, utilizou-se medidas de avaliação de qualidade do *ranking* de imagens sugeridas e a medida F, média harmônica entre precisão e revocação.

1.1 Motivação

Como motivação, o Exemplo 1 é uma típica situação na qual a percepção visual pode ser aplicada em sistemas de recomendação.

Exemplo 1 *Imagine uma pintura mostrando duas cenas, uma descrevendo um gato e outra descrevendo uma criança (Figura 1(a)). Suponha dois grupos de usuários U_{gato} e U_{crianca} . Suponha que os usuários de U_{gato} focaram sua atenção no gato e os de U_{crianca} focaram sua atenção na criança (Figura 1(b)).*

Em U_{gato} existem usuários que gostaram do quadro e usuários que não gostaram (Figura 1(b)). O mesmo ocorre com os usuários de U_{crianca} . As razões pelas quais alguns usuários de U_{gato} e alguns usuários de U_{crianca} gostaram do quadro são distintas, uma vez que eles observam a imagem de forma diferente.

Agora suponha que o novo usuário u^ tenha muito em comum com os usuários de U_{crianca} no que se refere à maneira como eles percebem uma pintura (Figura 2(a)). Assim,*

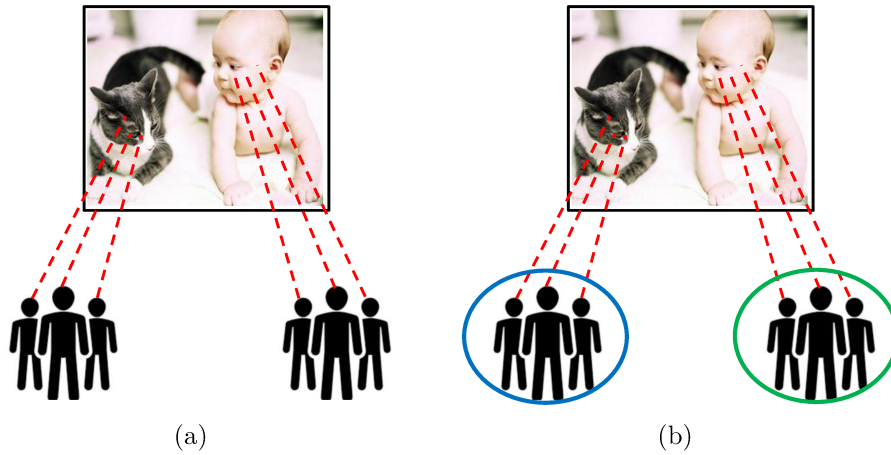


Figura 1 – Pintura sendo observada por dois grupos de pessoas.

é razoável se pensar em utilizar os usuários de $U_{crianca}$ no processo de recomendação de pinturas para o novo usuário u^ (Figura 2(b)).*

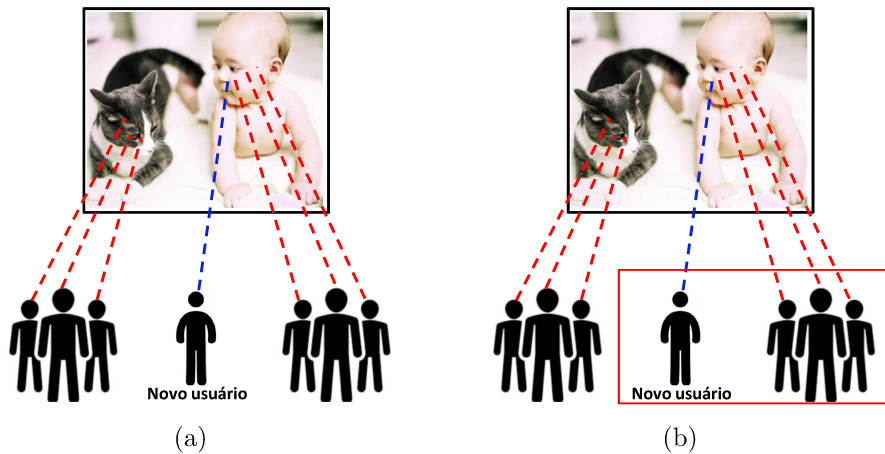


Figura 2 – Um novo usuário entra no sistema.

Nos sistemas de recomendação são utilizados dados convencionais (KUO et al., 2005; CHEN; HE, 2009; DAVIDSON et al., 2010; SANCHEZ et al., 2012). Os dados convencionais podem ser divididos em dados de usuários, dados de itens e dados de preferências. Dados de usuários são aqueles que referentes às características dos usuários como por exemplo idade, gênero e tipo de profissão (CHEN; HE, 2009). Dados de itens são aqueles cujos atributos (metadados) dos itens são levados em consideração. Por exemplo, no contexto de filmes, os dados de itens são atores, diretores, título e gênero (KUO et al., 2005; LEMOS et al., 2012; SANCHEZ et al., 2012). Dados de preferência, por sua vez, são atribuídos por meio das ações de usuários sobre objetos como, por exemplo, a indicação de nota por usuários em determinados objetos (DAVIDSON et al., 2010; WIDISINGHE et al., 2010; XU; JIANG; LAU, 2010; SUGANO et al., 2014).

Os dados não convencionais propostos na literatura são dados de contexto e dados de percepção visual (XU; JIANG; LAU, 2010; LIU; DOLAN; PEDERSEN, 2010; WIDISINGHE et al., 2010; UMEMOTO et al., 2012; SUGANO et al., 2014). Os dados de contexto são referentes ao contexto em que o usuário se encontra como, por exemplo, localização, situação que o usuário se encontra, data e hora. Um sistema de recomendação de pontos turísticos que utiliza dados de contexto, por exemplo, pode recomendar para um usuário que está em lua de mel, programas românticos.

Os dados de percepção visual são obtidos de acordo com as fixações oculares dos usuários em, por exemplo, imagens e páginas *web*. Tais fixações são utilizadas para definir preferências de usuários com base no que mais atrai a atenção deles. Por exemplo, um sistema de recomendação de roupas que utiliza dados de percepção visual pode inferir que um determinado usuário prefere roupas curtas pela sua fixação ocular na região das pernas de manequins em imagens de roupas e pela atribuição de notas boas para tais imagens. Então, o sistema de recomendação pode sugerir para esse usuário, por exemplo, bermudas, saias e vestidos curtos.

O uso dos dados não convencionais (contexto e percepção visual) foi proposto com a finalidade de melhorar a acurácia dos sistemas de recomendação e satisfazer as necessidades dos usuários. Assim, neste trabalho é proposta uma extensão do sistema de recomendação existente *PrefRec* (OLIVEIRA; AMO, 2013; AMO; OLIVEIRA, 2014), que faz uso de dados convencionais, combinando com dados de percepção visual. A Figura 3 ilustra a dinâmica dos dados utilizados no novo sistema - *VP-PrefRec*.

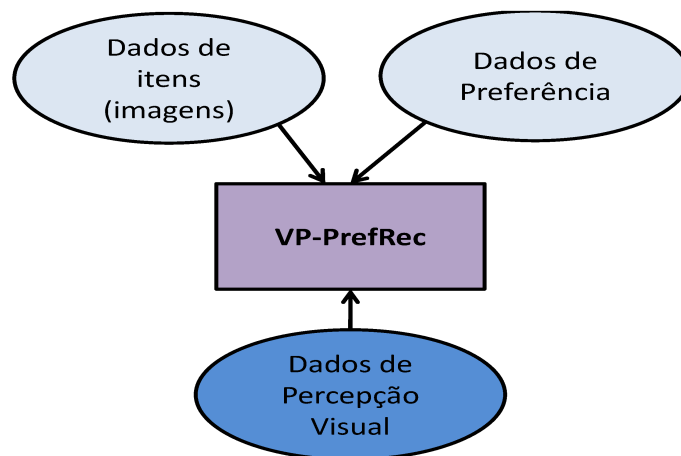


Figura 3 – Tipos de dados de entrada do *VP-PrefRec*.

O *PrefRec* foi desenvolvido por pesquisadores do mesmo grupo de pesquisa no qual este trabalho está incluído. Trata-se de um sistema de recomendação híbrido que faz uso de técnicas de mineração e agregação de preferências de usuários. A ideia original do *PrefRec* é que usuários que avaliaram um conjunto de itens de maneira semelhante apresentam perfis similares e podem fazer parte de um mesmo grupo. Assim, com base

nos dados dos itens avaliados por cada grupo, é possível prever preferências e usá-las na recomendação.

Como mencionado anteriormente, o uso de dados não convencionais pode melhorar a acurácia de sistemas de recomendação, e visto que o sistema *PrefRec* utiliza somente dados convencionais, a sua extensão, através de dados de percepção visual, pode ser vantajosa.

1.2 Objetivos

Para concretizar a proposta do uso da informação da percepção visual na recomendação de imagens, os principais objetivos deste trabalho são:

1. Propor um método para a representação da percepção visual de um usuário e de sua similaridade em relação a outro usuário;
2. Implementar o método proposto através da extensão do sistema *PrefRec* para utilização de dados de percepção visual em conjunto com dados de preferência e de itens na tarefa de recomendação;
3. Criar uma base de dados no domínio de pinturas, contendo dados de preferências e percepção visual dos usuários para a validação da proposta;
4. Aplicar o método proposto sobre um sistema de recomendação social, fornecendo uma rede de percepção visual, ao invés de uma rede social, para recomendação no contexto de usuários *cold-start*;

1.3 Hipótese

Esse trabalho é direcionado pelas seguintes hipóteses:

- H1: A recomendação de imagens com dados de itens e de preferências contextuais para um usuário u tem melhor qualidade quando são considerados apenas os usuários que têm percepções visuais similares às de u , em um cenário em que u já expressou algumas preferências.
- H2: É possível melhorar a qualidade das recomendações de imagens para um novo usuário u (u é *cold-start*), concentrando-se apenas em usuários que visualmente percebam as imagens de forma similar.

1.4 Organização da Dissertação

Esta dissertação está organizada em seis capítulos, como mostrado a seguir:

- ❑ Capítulo 1 - Apresenta a introdução, motivação e objetivos deste trabalho;
- ❑ Capítulo 2 - Descreve os trabalhos relacionados com essa dissertação e exibe uma tabela comparativa de todos os trabalhos citados;
- ❑ Capítulo 3 - Introduz os principais conceitos para um bom entendimento deste trabalho;
- ❑ Capítulo 4 - Apresenta todos os métodos proposto e a base de dados criada;
- ❑ Capítulo 5 - Discute os resultados dos experimentos realizados para validar as hipóteses desta dissertação;
- ❑ Capítulo 6 - Conclui esta dissertação e as principais contribuições deste trabalho, além de sugerir alguns possíveis trabalhos futuros.

Fundamentos Teóricos

2.1 Sistema de Recomendação

Desde meados dos anos 1990, quando começou os estudos sobre a filtragem colaborativa (RESNICK et al., 1994), o tema sistemas de recomendação se manifestou como uma interessante área de pesquisa independente (ADOMAVICIUS; TUZHILIN, 2005).

Os sistemas de recomendação são sistemas que ajudam os usuários a descobrir itens de seu interesse dentre as diversas opções existentes, evitando a sobrecarga de informação. Segundo (HERLOCKER et al., 2004), são três as principais etapas que compreende um sistema de recomendação, sendo elas: os dados de entrada dos clientes adquiridos através da preferência dos clientes, cálculo da recomendação por meio de técnicas apropriadas e a apresentação dos resultados obtidos com a recomendação para os clientes. Sistemas de recomendações são utilizados para fazer diversos tipos de recomendação como, por exemplo, de livros (LINDEN; SMITH; YORK, 2003), vídeos (DAVIDSON et al., 2010), páginas da *web* (BALABANOVIĆ; SHOHAM, 1997), imagens (SANCHEZ et al., 2012), músicas (KUO et al., 2005) e notícias (LIU; DOLAN; PEDERSEN, 2010).

Nos sistemas de recomendações, os interesses dos usuários geralmente são medidos por meio de classificações que indicam as preferências dos usuários. Por exemplo, se o usuário assistiu ao filme “Um amor pra recordar” e deu nota 2 (de 5) e assistiu também o filme “Karate Kid” e deu nota 5, quer dizer que ele prefere o filme “Karate Kid” ao filme “Um amor pra Recordar”. Outras formas de mensurar o interesse dos usuários é pela opção *like* de curtir ou não curti que muitos *websites* disponibilizam, o tempo que os usuários ficam em uma determinada página *web*, a quantidade de cliques feitos, os comentários prós e contras sobre um produto ou *website*, entre outras maneiras.

Conforme (ADOMAVICIUS; TUZHILIN, 2005), têm sido realizadas diversas pesquisas nessa área, uma vez que consiste de uma área de investigação com muitos problemas a serem solucionados, além de existirem diversos aplicativos que auxiliam os usuários a enfrentar a sobrecarga de informações, proporcionando recomendações personalizadas de serviços e conteúdos para os usuários. De acordo com (LINDEN; SMITH; YORK, 2003)

um bom sistema de recomendação para varejistas influentes como a Amazon.com, por exemplo, é aquele que possui um algoritmo que consiga gerar recomendações *on-line* com o mínimo de tempo de processamento possível, além de se adaptar rapidamente as oscilações nos dados de um usuário para que assim consigam fazer recomendações que satisfaçam todos os usuários, sem que haja dependência com número de avaliações e compras.

As técnicas de recomendação mais populares e que têm sido utilizadas por muitos pesquisadores são: a filtragem colaborativa, baseada em conteúdo e abordagem híbrida (HERLOCKER et al., 2004). Os detalhes sobre cada uma destas técnicas serão explorados nas três próximas subseções.

2.1.1 Técnica de Filtragem Colaborativa

A filtragem colaborativa é uma das técnicas que mais obteve sucesso nos sistemas de recomendação. Durante a última década, essa técnica vem sendo aperfeiçoada, assim, existem diversos algoritmos que utilizam essa técnica para geração de recomendações (HERLOCKER et al., 2004). Os sistemas de filtragem colaborativa procuram fazer predições de itens que sejam úteis para um certo usuário com base nos itens que tenham sido previamente avaliados por outros usuários (ADOMAVICIUS; TUZHILIN, 2005).

Nos sistemas colaborativos, as representações dos usuários são feitas de acordo com os itens que tenham sido classificados ou adquiridos. Por exemplo, em um sistema de recomendação de uma grande loja de eletrônicos, cada usuário é representado por um vetor de características onde são armazenados suas transações. Geralmente esse vetor é formado por 0 ou 1, sendo que 1 indica que um certo item foi comprado e 0 mostra que não houve transação de compra. Com isso, é criada uma matriz composta de todos os vetores que representam os usuários (HUANG; CHEN; ZENG, 2004).

Candillier, Meyer e Boullé (2007) apresentam três abordagens gerais para a filtragem colaborativa. A primeira é a abordagem baseada no usuário que associa um conjunto de vizinhos mais próximos para cada usuário, e em seguida, através das classificações, é feita a predição de um usuário sobre um item empregando as classificações dos vizinhos mais próximos desse item. A segunda consiste na abordagem baseada em modelos, especialmente a baseadas em *clusters*, onde é feito o agrupamento dos usuários, e depois prediz a classificação de um usuário sobre um item usando as classificações dos usuários de seu grupo. E por fim, a abordagem baseada em itens, que associa um conjunto de vizinhos mais próximos para cada item, e posteriormente prevê a classificação de um usuário sobre um item por meio das classificações do usuário sobre os vizinhos mais próximos do item analisado.

Apesar de ser uma técnica bem promissora na área de sistema de recomendação, a filtragem colaborativa tem alguns desafios importantes para serem superados. O principal problema é a dispersão dos dados causada pelo fato das matrizes de iteração serem bas-

tantes esparsas. Isso ocorre porque nos grandes *websites*, tanto o número de itens como o número de consumidores são muito grandes, isso faz com que nem todos os produtos sejam comprados ou avaliados. Outro problema bem conhecido é o *cold start de item* que faz com que o sistema não consiga fazer recomendações de novos itens do sistema por causa da falta de avaliações ou compras feitas com este item (SCHEIN et al., 2002; HUANG; CHEN; ZENG, 2004; LAM et al., 2008; QIU et al., 2011; SHANI; GUNAWARDANA, 2011).

2.1.2 Técnica Baseada em Conteúdo

A abordagem baseada em conteúdo teve início na filtragem e recuperação de informações (WEI; HUANG; FU, 2007). De acordo com (HERLOCKER, 2000) a origem do nome dessa técnica se deu pelo fato do *software* gerado conseguir selecionar itens de acordo com o conteúdo dos itens analisados. Os sistemas de recomendação baseados em conteúdo são capazes de extrair os detalhes do conteúdo de cada item, com isso conseguem comparar as informações dos usuários com os conteúdos dos itens e determinar se um item é relevante ou não de acordo com a necessidade de cada usuário. As informações que formam o perfil de cada usuário podem ser obtidas através das avaliações que os usuários fazem sobre itens. Essas informações podem ser passadas através do próprio usuário ou através do conteúdo dos itens.

Conforme (WEI; HUANG; FU, 2007) os itens recomendados costumam ser descritos com pesos e palavras-chaves. A análise e agrupamento do conteúdo das características textuais dos itens são feitas por método de agrupamento ou por funções de vizinhos mais próximos. Em seguida, os conteúdos são recomendados com base nas preferências dos usuários e características dos itens. Geralmente, os itens recomendados pela recomendação baseada em conteúdo sugerem informações textuais como documentos e redes de notícias.

Os sistemas de recomendação baseado em conteúdo possuem como uma de suas principais vantagens, a recomendação de itens novos mesmo que esse itens ainda não tenham sido avaliados por outros usuários (BARRILERO et al., 2011). Com isso o problema do *cold start* do item da filtragem colaborativa é solucionado com essa técnica. Contudo, essa abordagem possui alguns problemas (LOPS; GEMMIS; SEMERARO, 2011). Um dos problemas é a super especialização. A recomendação baseada em conteúdo não é capaz de fazer recomendações de novidades. Isso acontece porque o sistema só recomenda itens que possuem avaliação alta quando comparada com o perfil do usuário, portanto itens que não sejam semelhantes aos já classificados, dificilmente serão recomendados. Além disso, os sistemas baseados em conteúdos sofrem com o problema de *cold start* do usuário, que faz com que os sistemas não consigam fazer recomendações para novos usuários (LAM et al., 2008; LOPS; GEMMIS; SEMERARO, 2011; BERNARDI et al., 2015). Isso acontece por causa da falta de avaliações dos novos usuários, assim, os sistemas têm dificuldades em prever as preferências dos usuários.

2.1.3 Abordagem Híbrida

A abordagem de recomendação híbrida combina duas ou mais técnicas de recomendação com a finalidade de conseguir um melhor desempenho dos sistemas com a mínima desvantagem para os usuários (BURKE, 2002). Muitos sistemas de recomendação estão utilizando suas próprias abordagens híbridas (BALABANOVIĆ; SHOHAM, 1997; BURKE, 2005; BURKE, 2007; WOERNDL; SCHUELLER; WOJTECH, 2007; OLIVEIRA; AMO, 2013; AMO; OLIVEIRA, 2014). Isso é feito combinando técnicas da filtragem colaborativa com as baseadas em conteúdo. Com isso, é possível minimizar o problema do *cold start do item* causado pela filtragem colaborativa e o *cold start do usuário* provocado pela baseada em conteúdo (CHOI; JEONG; JEONG, 2010).

Segundo (ADOMAVICIUS; TUZHILIN, 2005), em um sistema de recomendação híbrido existem várias maneiras de combinar a filtragem colaborativa com a baseada em conteúdo. Uma forma é implementando cada técnica separadamente e depois combinando as suas predições. Outra maneira é introduzindo certas características da técnica baseada em conteúdo na filtragem colaborativa ou então, incorporando na abordagem baseada em conteúdo algumas características da filtragem colaborativa. Pode-se também construir um modelo que unifique as duas técnicas, ou seja, um modelo que contenha tanto características da filtragem colaborativa quanto da baseada em conteúdo.

2.2 Percepção visual

Uma área de pesquisa que tem atraído a atenção dos pesquisadores da neurociência computacional é como desenvolver modelos computacionais que consigam representar a atenção de uma pessoa em uma dada imagem (ITTI; KOCH, 2001). Uma imagem pode gerar diferentes modelos de percepções para diferentes usuários. De acordo com (WAN; JIN; YUE, 2009), estudos realizados na psicologia visual destacam que existe partes de uma determinada imagem que podem atrair mais a atenção de uma pessoa do que de outra ao observar a mesma imagem. Quando se analisa uma imagem, as pessoas tendem primeiramente a observar as regiões que despertam impulsos, como por exemplo, regiões com cores vivas. Logo após, elas costumam prestar atenção em regiões que chamam sua atenção, como por exemplo, um detalhe que lhes parece interessante.

Koch e Ullman (1987) foram uns dos pioneiros no estudo da atenção visual. Eles implementaram um framework para determinar as regiões de uma imagem através de mapas de saliência. Esses mapas representam características importantes como cor, orientação e direção de um movimento. Desde então diversos modelos foram propostos, dentre eles (TREISMAN; GELADE, 1980), (SANDON, 1990), (AHMAD, 1991).

2.3 Extração de características de imagem

Em conformidade com (POST et al., 2003), as técnicas de processamento de imagens tiveram início na análise de dados de imagens em 2D e 3D. Essas imagens costumam ser representadas como valores escalares em tons de cinza em uma grade retangular.

Segundo (YAN, 2011), sistemas que utilizam a extração de características de imagens são sistemas de busca, navegação e recuperação de imagens de banco de dados volumosos de imagens digitais. Um método comum para recuperação de imagens é através de textos que descrevem as imagens, porém este método possui alguns inconvenientes. Normalmente uma mesma imagem pode gerar diferentes descrições entre dois ou mais usuários. Dessa forma, é difícil descrever essa imagem através de um rótulo e responder com precisão a busca de um usuário. Além disso, a maioria das imagens possuem muitos detalhes e significados amplos. Com isso, é trabalhoso especificar uma imagem por meio de palavras chaves ou comentários. Uma forma de contornar esses problemas é através da extração de características da própria imagem, pois a extração é objetiva e todo o processo pode ser realizado automaticamente.

Neste trabalho, dividimos o processamento digital de imagens em duas áreas de pesquisa. A primeira área é focada nos sistemas de recuperação de imagens conhecidos como CBIR (Content-Based Image Retrieval). Esses sistemas utilizam técnicas de visão computacional para resolver os problemas relacionados a recuperação de imagem, ou seja, o problema de procurar imagens em grandes bases de dados. Por exemplo, o usuário fornece uma imagem de base para o sistema e ele é responsável por retornar as imagens similares aquelas que o usuário mostrou (BRODLEY et al., 1999). A segunda área de pesquisa é constituída pelos métodos utilizados para extração de características de imagens que ajudam na criação dos sistemas de recomendação de imagens. O intuito é descobrir através da extração de características de imagens, juntamente com as percepções visuais dos usuários, o que atrai a atenção das pessoas fazendo com que elas gostem ou não de uma imagem. A seguir serão descritas as principais características que frequentemente são extraídas das imagens são cor, forma e textura:

1. Cor: A cor é um importante descritor que ajuda na identificação e extração de um objeto numa imagem. Os dois principais modelos de cores existentes são o RGB e o HSI. De acordo com (GONZALEZ; WOODS, 2010), a finalidade dos modelos de cores é simplificar a classificação das cores em alguma forma padronizada.

□ Modelo de cor *RGB* (*Red*, *Green*, *Blue*): esse modelo é baseado em um sistema de coordenadas cartesianas. O cubo apresentado na Figura 4 representa o subespaço de cores. Conforme (GONZALEZ; WOODS, 2010), é possível perceber que os valores RGB primários estão em três vértices; as cores secundárias magenta, ciano e amarela estão em outros três vértices; o preto encontra-se na

origem enquanto o branco está no vértice mais afastado da origem. Além disso, a escala de cinza vai do preto até o branco através do segmento de reta que une esses dois pontos. O modelo RGB é utilizado para monitores coloridos e várias classes de câmeras de vídeo em cores.

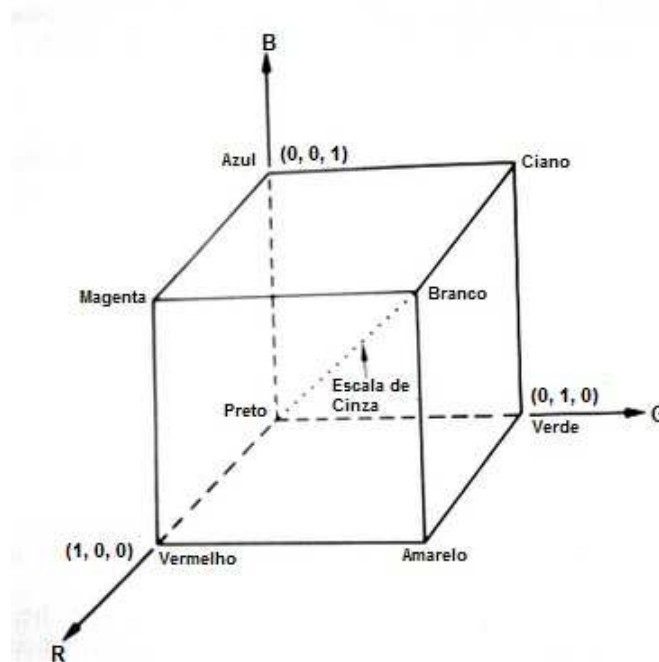


Figura 4 – Representação de um cubo RGB. Fonte: (GONZALEZ; WOODS, 2010).

- Modelo de cor *HSI* (*Hue, Saturation, Intensity*). As cores utilizadas nesse modelo equivale a forma como os seres humanos descrevem e compreendem as cores. Segundo (GONZALEZ; WOODS, 2010) o modelo HSI separa o componente intensidade das informações de cores (matiz e saturação) em uma imagem colorida. Assim, esse modelo é considerado uma boa ferramenta para o desenvolvimento de algoritmos que são baseados em descritores de cores que são evidentes e naturais para os seres humanos.

Uma maneira de representar as cores é por meio de histogramas. O histograma de cor pode ser adquirido por meio da discretização das cores de uma imagem e pela contagem da quantidade de vezes que cada cor aparece na matriz da imagem (SWAIN; BALLARD, 1991). Eles não variam de acordo com a rotação e translação em torno do eixo da visão, ou seja, a movimentação de um certo objeto em uma imagem não altera o histograma dessa imagem. A Figura 5 exibe um exemplo de uma imagem com o seu histograma.

2. Forma: A forma é uma importante característica que pode ser extraída de uma imagem. Ela ajuda na identificação de objetos numa imagem através dos contornos formados nas regiões de interesses. Um grande desafio na extração de formas é

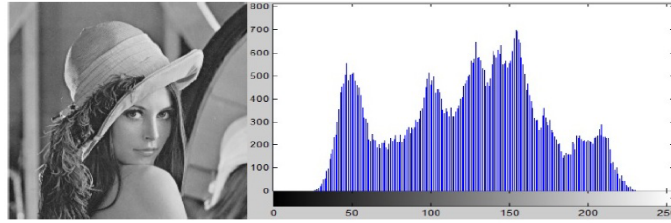


Figura 5 – Exemplo de uma imagem e seu histograma. Fonte: (GONZALEZ; WOODS, 2010).

o processo de segmentação das imagens que dependem bastante do domínio da imagem. Os métodos de segmentação dividem uma imagem em partes ou objetos que a constitui. Um bom processo de segmentação de imagens amplia as chances de sucessos na resolução de problemas que necessitam que os objetos sejam identificados individualmente (GONZALEZ; WOODS, 2010).

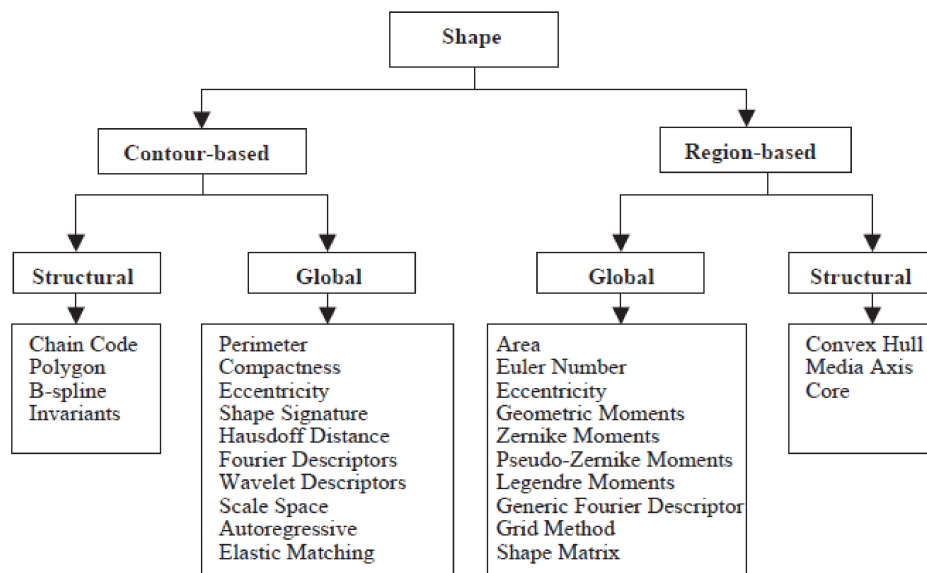


Figura 6 – Hierarquia das técnicas de extração de formas. Fonte: (ZHANG; LU, 2004).

Os algoritmos de segmentação para imagens monocromáticas frequentemente se baseiam em discontinuidades ou similaridades (GONZALEZ; WOODS, 2010). Esses são dois grupos básicos tem relação com as especificidades dos valores de intensidade. O primeiro grupo usa a hipótese que as fronteiras das regiões são bastante diferentes uma das outras e também em relação ao fundo da imagem, o que permite a detecção de limites com base nas discontinuidades locais em intensidade. A abordagem dominante nesse grupo é a segmentação baseada em bordas. O segundo grupo baseia-se na divisão de uma imagem em regiões que sejam similares conforme alguns critérios pré-definidos. Um exemplo de abordagem utilizada nesse grupo é a segmentação baseada na região. A Figura 6 mostra a hierarquia destes dois grupos

(ZHANG; LU, 2004).

3. Textura: A textura auxilia na compreensão correta de uma cena, fornecendo medidas de propriedades como suavidade, rugosidade e regularidade. Segundo (REED; DUBUF, 1993) tem sido proposto diferentes maneiras de segmentação da textura. Geralmente a segmentação é adquirida pela seleção de características de textura, por subprocessos independentes de extração de características ou por redução de características seguido por um algoritmo de segmentação. Em geral, os métodos de extração de características de texturas são classificados em três grupos: baseados em características, baseados em modelo e estrutural (REED; DUBUF, 1993). Os métodos baseados em características faz a seleção de algumas características e regiões onde essas características não variam muito. Os métodos baseados em modelo supõem processos subjacentes para texturas e segmentos utilizando alguns parâmetros desses processos. Já os métodos estruturais buscam dividir as imagens usando como hipótese que as texturas na imagem possuem elementos primitivos identificáveis, organizado conforme as regras de colocação. Os métodos baseados em modelo poderiam ser apontados como uma subclasse dos métodos baseados em características, uma vez que os parâmetros do modelo são usados como características de textura.

2.4 PrefRec

Em (OLIVEIRA; AMO, 2013; AMO; OLIVEIRA, 2014) é proposto o *PrefRec*, um sistema de recomendação híbrido que utiliza técnicas de mineração e agregação de preferências com o intuito de fazer recomendações acuradas.

O sistema adota os conceitos gerais da recomendação baseada em conteúdo e da filtragem colaborativa. Os usuários que possuem gostos similares são agrupados. Em seguida, é definido um gosto consensual para cada um dos grupos criados. Esse gosto consensual é submetido a um minerador de preferências contextuais. Como resultado do minerador é gerado um conjunto de regras de preferências contextuais. Essas regras serão usadas para prever preferências de usuários que se encaixam em um determinado grupo.

A Figura 7 apresenta o esquema geral do *PrefRec*. De acordo com (AMO; OLIVEIRA, 2014), o ***PrefRec*** é composto por cinco módulos: os 4 primeiros módulos trabalham de modo *offline* e correspondem aos módulos de construção do modelo de recomendação. O último módulo é executado *online* e é responsável pela recomendação. A seguir uma breve descrição dos cinco módulos.

1. Módulo de Representação da preferência: tem por objetivo pré-processar as preferências de cada usuário por uma determinada amostra. Diferente da representação dos usuários normalmente utilizada em sistemas de recomendação que representam

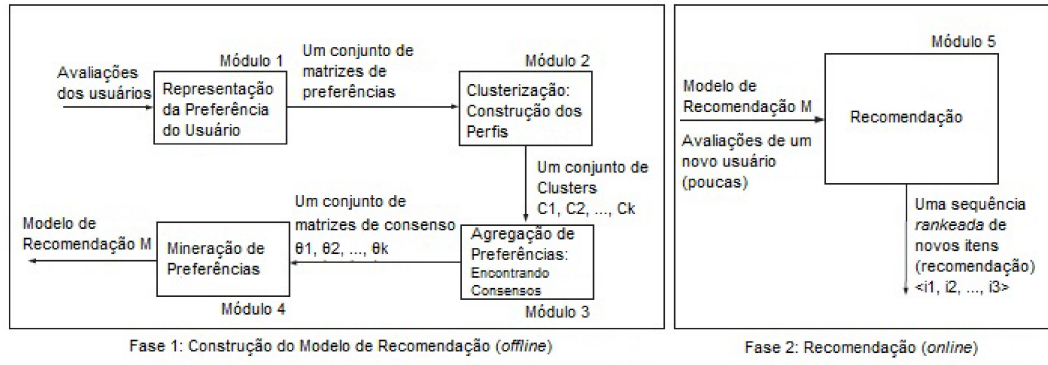


Figura 7 – Esquema geral do *PrefRec*. Adaptado de: (AMO; OLIVEIRA, 2014).

cada usuário por um vetor de notas (Figura 8 (a)), no PrefRec é utilizado o conceito de matriz de preferências (*Mpref*) como uma nova maneira de representar os usuários em um sistema de recomendação. Uma *Mpref* é uma matriz com dimensão $n \times n$, sendo n a quantidade de itens do sistema. Cada posição $(a; b)$ da *Mpref* de um determinado usuário contém um valor entre 0 e 1 que representa o quanto o usuário prefere o item a ao item b , isto é, grau de preferência do usuário ao item a em relação ao item b . A Figura 8 é um exemplo de uma *Mpref*.

Item	Avaliacao
i_1	5
i_2	3
i_3	*
i_4	1

(a)

	i_1	i_2	i_3	i_4
i_1	0.50	0.63	*	0.83
i_2	0.37	0.50	*	0.75
i_3	*	*	0.50	*
i_4	0.17	0.25	*	0.50

(b)

Figura 8 – Exemplo de uma *Mpref* de um usuário: (a) Avaliações de um usuário. (b) Matriz retornada pelo Módulo 1, correspondendo as preferências do usuário. Os (*) nas tabelas significam que determinado item não foi avaliado pelo usuário.

A saída deste módulo é um conjunto de matrizes, uma para cada usuário.

2. Módulo de Clusterização: neste módulo o usuário é agrupado de acordo com suas preferências. Os usuários são organizados em grupos a partir das suas respectivas *Mprefs* calculadas na etapa 1. Com isso, os usuários com gostos semelhantes ficam no mesmo grupo C_i . Essa etapa resulta diretamente na escalabilidade do Sistema de Recomendação em relação ao momento de recomendação. O usuário ativo será comparado a quantidade de perfis existentes e não a todos os usuários do conjunto, como é feito tradicionalmente. A quantidade de perfis é bem menor do que a quantidade de usuários. A saída deste módulo é um conjunto de *clusters* de preferências,

onde cada grupo é constituído por um conjunto de matrizes de preferência que são similares indicando a similaridade de gosto dos usuários do grupo.

3. Módulo de Agregação de Preferências: o objetivo deste módulo é a partir de um grupo de usuário com gostos similares, produzir um consenso de preferência. É obtido uma M_{pref} consensual para cada grupo a partir da agregação das M_{pref} s dos seus respectivos usuários. A M_{pref} consensual corresponde ao consenso C do perfil de preferência de cada grupo. A saída deste módulo é um conjunto de matrizes com as preferências consensuais de cada *cluster* de preferência.
4. Módulo de Mineração de Preferência: este módulo é responsável pela construção de um modelo de preferência M para cada *cluster*. Com isso, um modelo de preferência é capaz de prever, dados dois itens, qual o preferido. O modelo de preferências M traduz a agregação das preferências individuais de um grupo de usuários que possuem preferências semelhantes. Esse modelo é utilizado para calcular as recomendações ao usuário ativo a partir do momento que se identifica sua similaridade com algum grupo. Desta forma, o usuário ativo receberá recomendações de itens que o grupo, ao qual ele é semelhante. A saída deste módulo é o modelo de recomendação M .
5. Módulo de Recomendação: este módulo tem como objetivo fazer recomendações de novos itens para usuários a partir do modelo de recomendação M . Nessa etapa a recomendação é feita online, ao usuário ativo. Tem-se a disposição o modelo de recomendação contruído nas etapas anteriores, e dados do histórico de avaliações do usuário ativo. A saída deste módulo é uma sequência de novos itens rankeados que serão recomendados para o novo usuário.

Revisão Bibliográfica

Existem vários trabalhos sobre sistemas de recomendações. Trata-se de uma área de investigação com muitos problemas e aplicações práticas. Essas aplicações tem o intuito de ajudar os usuários a lidar com a sobrecarga de informações e proporcionar recomendações mais personalizadas.

A Figura 9 exibe um diagrama com os tipos de dados existentes, sendo que no nosso sistema foram utilizados dados de preferências, dados de itens e dados de percepção visual que estão representados pela cor azul.

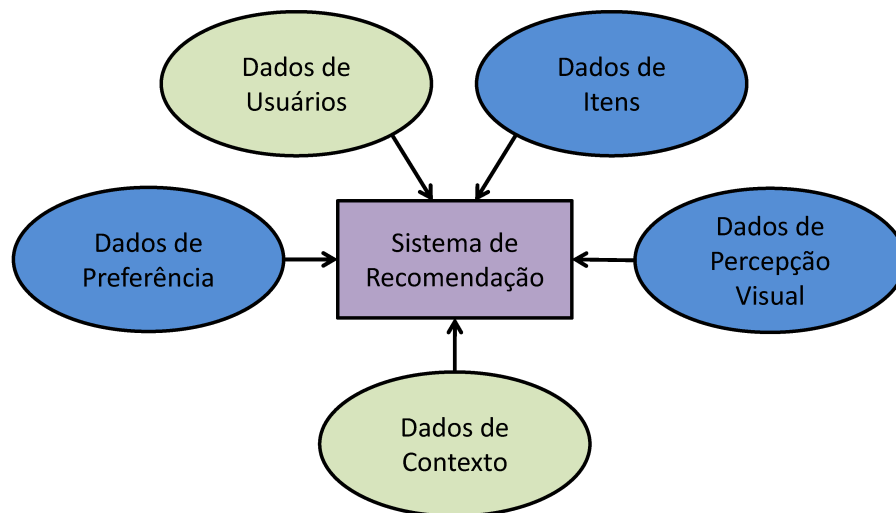


Figura 9 – Tipos de dados dos sistemas de recomendação.

De acordo com a Figura 9, os dados dos sistemas de recomendação podem ser categorizados em 5 grupos. São eles: baseados em dados de usuários, dados de itens, dados de preferências, dados de percepção visual e dados de contexto.

Sistemas de recomendação baseados em dados dos usuários são os que consideram as características de cada usuário na hora de fazer recomendações. Por exemplo, um sistema que recomenda produtos de uma loja deverá considerar os dados do usuário para poder sugerir algum item ao usuário. Caso seja um usuário do sexo feminino e com idade inferior

a 7 anos, o sistema poderá recomendar brinquedos, como boneca. Porém, se o usuário for do sexo masculino e também com idade inferior a 7 anos, o sistema deverá recomendar brinquedos relacionados a meninos, como um carrinho por exemplo. Supondo agora que um usuário tem a profissão de cozinheiro, o sistema fará recomendações de livros de culinárias ou materiais utilizados nesse tipo de profissão.

Sistemas de recomendação baseado nos dados de itens são aqueles nos quais os atributos do objeto são levados em consideração na hora de predizer as preferências dos usuários. Um exemplo desse tipo de sistema é um recomendador de filmes. Nesse tipo de sistema são consideradas características do item, como o título do filme, gênero, diretor e ator para o processo de recomendação. Se um usuário x gosta de filme de comédia com o ator *Jim Carrey*, por exemplo, então o sistema deverá recomendar filmes como “Os Pinguins do Papai”, “Debi e Lóide”, “O Golpista do Ano”, entre outros que são do gênero comédia e protagonizado pelo ator *Jim Carrey*. Supondo que seja um sistema recomendador de imagens, a extração das características das imagens tais como, cor, textura, forma também são consideradas como dados do objeto. Considere que é exibido ao usuário um conjunto de 10 imagens para ele indicar se gosta ou não de cada imagem. Se em cinco imagens predomina a cor azul e o usuário gostou delas, então quer dizer que a cor azul atrai a atenção desse determinado usuário fazendo com que o sistema recomende imagens que tenham essa cor.

Sistemas de recomendação baseado em dados de preferências são sistemas que consideram as preferências (gostos) dos usuários na hora de fazer recomendações. Os gostos do usuário podem ser obtidos através de um *feedback* explícito ou implícito (LEE; PARK; PARK, 2008; JAWAHEER; SZOMSZOR; KOSTKOVA, 2010; HU; KOREN; VOLINSKY, 2008). Exemplos de *feedback* explícitos são os likes de curtir ou não curtir que muitas páginas *web* usam, ou também podem ser notas variando de 1 a 5 sendo que a nota é proporcional ao gosto e, podem ser inferidos por meio de comentários que um usuário faz sobre algum produto. Exemplos de *feedbacks* implícitos são a quantidade de cliques feitos sobre determinada página ou objeto e o tempo que um usuário ficou assistindo um vídeo ou permaneceu em uma página.

Sistemas baseados em dados de percepção visual fazem recomendações considerando a percepção do usuário para identificar as preferências (SUGANO et al., 2014; UMEMOTO et al., 2012; XU; JIANG; LAU, 2010). Esse tipo de sistema atua de acordo com as fixações oculares dos usuários. Para auxiliar nas recomendações, esses sistemas geralmente associam notas junto com a percepção visual do usuário. Suponha uma nota variando de 1 a 5 e que temos uma imagem de vários animais incluindo um cachorro. Se o usuário atribui a nota 1 ao fixar o olhar muito tempo no cachorro, pode-se entender que aquele determinado usuário não gosta muito de cachorro. O contrário também é verdadeiro, caso ele fixe muito o olhar em um cachorro e atribui nota 5 (nota máxima) quer dizer que aquele animal chamou a atenção do usuário fazendo com o que ele gostasse mais daquela

imagem.

Sistemas de recomendação baseado em dados de contexto levam em consideração além de atributos, o contexto em que o usuário se encontra. Os atributos que podem ser considerados nesses sistemas são: localização, data, hora, situação que a pessoa se encontra. Por exemplo, se um usuário está de férias com a família e ainda está de dia, o sistema deverá recomendar programas que possam ser feitos em família como ir a praia ou passear em um zoológico. Se esse mesmo usuário está de férias com sua esposa e é de noite, então o sistema deverá recomendar programas que possam ser feitos a dois, como um jantar romântico.

A seguir serão apresentados os trabalhos correlatos mais recentes que tem por objetivo a recomendação de itens. Além disso, também serão apresentados os trabalhos que utilizam dados de percepção visual coletados por meio de um rastreador ocular.

3.1 Sistemas de Recomendação

No trabalho de (CHEN; HE, 2009) é descrito um método para minimizar um dos principais desafios dos sistemas de recomendação que é o problema do *cold start*, apresentado em maiores detalhes na Seção 2.1. É proposto um método que utiliza a técnica de filtragem colaborativa baseada num vetor de atributos demográficos. Esse método é centrado no cálculo da similaridade de novos usuários. O cálculo é feito a partir dos atributos demográficos dos usuários, como por exemplo: idade, sexo, área de atuação e ocupação. A abordagem utilizada é baseada na seguinte suposição: caso duas pessoas possuam ocupação, idade, renda, sexo e outros dados semelhantes, então os interesses delas podem ser por itens semelhantes. Os resultados obtidos demonstram que o método proposto pode melhorar a qualidade dos sistemas de recomendação. O algoritmo apresentado depende dos dados cadastrais dos usuários que precisam estar todos preenchidos.

Sanchez et al. (2012) propõem um sistema de recomendação híbrido que considera a estética e as características formais das imagens para fazer recomendações de imagens em redes sociais móveis, com alto grau de adaptação a qualquer tipo de usuário e interface. O sistema é composto por dois processos. O primeiro é denominado processo de classificação e trabalha de modo *offline* para extrair as características das imagens e armazená-las em uma base de dados. O segundo processo, por sua vez, é denominado processo de recomendações, funciona de modo *online* e é responsável por gerar uma lista de recomendações para cada usuário. A saída do sistema é uma lista com imagens recomendadas.

Em (KUO et al., 2005) é proposto um sistema de recomendação genérico de músicas para filmes, baseado na emoção que as músicas proporcionam para os filmes. Considerando que as músicas exercem um papel fundamental na propagação de emoções nos filmes, o sistema proposto neste artigo extrai as características das músicas com o objetivo de encontrar uma relação entre essas características e as emoções. Exemplos comuns

de recomendações de músicas baseadas na emoção inclui as músicas tocadas em *shopping* para estimular a venda, música tocadas em casa para acomodar os habitantes, a musicoterapia.

Um sistema de recomendação híbrido de notícias personalizadas para o *Google News* foi implementado em (LIU; DOLAN; PEDERSEN, 2010). O sistema foi desenvolvido com o intuito de ajudar na localização de artigos de notícias que atendem aos interesses dos usuários. Os perfis dos usuários são construídos de acordo com o *feedback* implícito de cliques passados. Diante da necessidade de compreender como os interesses dos usuários variam de acordo com o tempo, foi realizado um estudo de *logs* de cliques do *Google News*. Com isso, foi descoberto que os interesses dos usuários não variam muito ao longo do tempo, porém seguem uma tendência global, uma vez que a distribuição dos cliques em geral refletem a notícias de grandes eventos. Além disso, os interesses por notícias diferem dependendo dos diferentes locais em que os usuários se encontram. Foi proposto um *framework* bayesiano que é capaz de descobrir o que realmente interessa aos usuários por meio da combinação das tendências atuais de notícias com o que o usuário se sente atraído para prever notícias de interesse dos usuários. Por fim, o método híbrido melhorou significativamente a qualidade das recomendações de notícias.

Davidson et al. (2010) descrevem um sistema de recomendação de vídeo utilizado no *YouTube*. O propósito do sistema é proporcionar recomendações personalizadas aos interesses dos usuários. As recomendações são atualizadas frequentemente e são baseadas nas atividades recentes dos usuários. Desse modo, o sistema consegue manter os usuários entretidos com o *site*. Além disso, as recomendações do *Youtube* são feitas preservando a privacidade dos usuários e fornecendo o controle explícito sobre dados de usuários personalizados. O conjunto de vídeos recomendados pelo sistema é gerado utilizando a atividade pessoal de cada usuário, como por exemplo os *likes* de curtir ou não curtir. Esse conjunto de dados é estendido com um gráfico baseado na co-visitação de vídeos. Assim sendo, usando uma variedade de sinais de relevância e diversidades, os conjuntos de vídeos são classificados. O *Youtube* faz uso desse sistema de recomendação há mais de um ano e até então tem obtido sucesso nas recomendações (DAVIDSON et al., 2010).

O sistema *MMedia2U*, descrito em (LE MOS et al., 2012) faz recomendações de fotos por meio de dados de contexto. Esse tipo de recomendação considera as características inerentes a cada usuário e item, e também a peculiaridade da situação atual. Tentando evitar o problema *cold start* da filtragem colaborativa, esse sistema usa um método de reconhecimento baseado em conhecimento. As imagens recomendadas pelo sistema servem como referências de alguma atividade ou novos lugares que podem ser explorados perto da localidade do usuário ou para oferecer uma nova perspectiva da situação atual.

Widisinghe et al. (2010) apresentam o *picSEEK*, um sistema de recomendação de imagens que considera dados de contexto do usuário. O sistema consegue identificar preferências de imagens para um usuário em qualquer contexto, por meio de experiências

passadas como, por exemplo, o *feedbacks* implícito de imagens selecionadas por usuários de contexto semelhantes e com pensamentos similares. Desse modo, Widinghe et al. (2010) exploram técnicas da filtragem colaborativa para um contexto de recomendação que usa diferentes tipos de domínios de contexto. É descrito dois algoritmos para recomendação de imagens. O primeiro algoritmo baseia-se em rótulos de imagens, enquanto o segundo é estruturado em relação contexto em que duas ou mais imagens foram selecionadas. Nessa aplicação o cliente pode escolher um desses algoritmos, dependendo de sua necessidade.

3.2 Sistemas de Recomendação que utilizam rastreadores de olhar

Uma abordagem orientada a dados para estimar a preferência de imagens por meio dos movimentos oculares foi proposto em Sugano et al. (2014). Como dados de entrada são armazenados as fixações e os movimentos oculares dos usuários quando estão comparando um par de imagens. A partir desses conjuntos de dados e métodos estatísticos são gerados os vetores de características. Em seguida, através de uma tarefa supervisionada, um algoritmo de *random forest* é treinado para estimar os rótulos de preferências desconhecidas das imagens.

Umemoto et al. (2012) propõem um sistema que tem o intuito de ajudar os usuários a encontrar de forma eficiente informações que são relevantes em páginas *web*. O trabalho foi dividido em duas fases. Na primeira é proposto um método para calcular os interesses dos usuários em páginas *Web* por meio dos seus movimentos oculares. Para validar esse método, eles partem do pressuposto que quanto mais o usuário olhou para um determinado termo numa página de busca, mais essa palavra está relacionada com o que ele procura. Na segunda fase é proposto e implementado um sistema cujo objetivo é ajudar os usuários a encontrar com mais facilidade informações que são do interesse dos usuários nos navegadores *web*. Esse sistema utiliza o movimento ocular dos usuários, que é obtido por meio de um rastreador ocular conhecido como *Tobi*. Uma das grandes dificuldades encontradas pelos autores é quando o usuário acessa uma página que possuem animações ou páginas com muitas imagens. Esses recursos tendem a chamar mais atenção dos usuários do que o texto da página o que prejudica o desempenho do sistema.

Xu, Jiang e Lau (2010) propõem um método que faz recomendação de páginas *web* personalizadas por meio da ordenação dos interesses de leitura dos usuários. O algoritmo proposto utiliza como dados de entrada a expressão facial dos usuários e sua posição ocular no decorrer de suas atividades de navegação em uma determinada página *web*. Para obter os movimentos oculares dos usuários é utilizada uma câmera juntamente com técnicas de visão computacional. Desse modo, é possível traçar os interesses de leituras de cada usuário por meio de *feedbacks* implícitos obtidos pelos comportamentos de navegação

na *web*. Com base nas informações adquiridas o algoritmo consegue reconhecer possíveis interesses dos usuários em outras páginas que o usuário ainda não tenha acessado no conjunto de resultados de uma busca na *web*. A qualidade do algoritmo desenvolvido no trabalho foi comparado com os navegadores de buscas comerciais e tem obtido bons resultados na recomendação de páginas *web* personalizadas.

A Tabela 1 sintetiza a comparação de todos os trabalhos relacionados estudados sobre sistemas de recomendação.


Tabela 1 – Comparação dos Trabalhos Correlatos

Trabalhos	Tipo de Entrada	Tipo de Saída	Tipo de Dados
(KUO et al., 2005)	Conjuntos de músicas e emoções geradas pelas músicas.	Sistema de recomendação genérico de músicas para filmes de cinema.	Dados de objetos
(CHEN; HE, 2009)	Dados demográficos.	Sistema de recomendação que utiliza a técnica de filtragem colaborativa baseada num vetor de atributos demográficos.	Dados de usuários
(DAVIDSON et al., 2010)	Conjunto gerado pela atividade pessoal de cada usuário no youtube, além de um gráfico baseado na co-visitação de vídeos.	Sistema de recomendação de vídeo utilizado no <i>YouTube</i> .	Dados de preferência
(WIDISINGHE et al., 2010)	<i>Feedback</i> implícito de imagens selecionadas por usuários em contextos semelhantes.	Recomendações de imagens considerando o contexto que os usuários se encontram.	Dados de preferência e dados de contexto
(LIU; DOLAN; PEDERSEN, 2010)	<i>Feedback</i> implícito de cliques passados dos usuários.	Sistema de recomendação híbrido de notícias personalizadas para o Google News.	Dados de preferência e dados de contexto
(XU; JIANG; LAU, 2010)	Expressão facial dos usuários, movimentos oculares e <i>feedback</i> implícito obtidos pelo comportamento de navegação do usuário.	Sistema de recomendação de páginas <i>web</i> por meio da ordenação dos interesses dos usuários.	Dados de preferência e dados de percepção visual
(SANCHEZ et al., 2012)	Conjunto de imagens.	Sistema de recomendação híbrido que considera a estética e as características formais das imagens para fazer recomendações de imagens em redes sociais móveis.	Dados de objetos
(LEMOES et al., 2012)	Contexto em que o usuário se encontra, além das características de cada usuário e item.	Recomendações de fotos através de um contexto consciente.	Dados de usuários, dados de objetos e dados de contexto
(UMEMOTO et al., 2012)	Páginas <i>web</i> , fixações e movimentos oculares.	Recomendações de informações relevantes em páginas <i>Web</i> .	Dados de percepção visual
(SUGANO et al., 2014)	Fixações e movimentos oculares dos usuários	Sistema de recomendação de imagens estimadas através dos movimentos oculares	Dados de preferência e dados de percepção visual
Trabalho proposto	Conjunto de imagens, fixações, movimentos oculares e notas representando as preferências dos usuários.	Sistema de recomendação de imagens baseado na percepção humana e nas preferências dos usuários.	Dados de objetos, dados de preferência e dados de percepção visual

Proposta

É proposto neste trabalho um método geral para calcular a similaridade de percepção visual entre os usuários, chamado *VP-similarity*. Esse método foi aplicado em um sistema de recomendação híbrido baseado na mineração de preferência em pares, *PrefRec* (OLIVEIRA; AMO, 2013). Denominamos a junção do *VP-similarity* ao *PrefRec* como *VP-PrecRec*.

A fim de tratar o problema do *cold start*, também foi proposto a inferência de uma rede de percepção visual dos usuários. Além disso, criou-se uma base de dados de percepção visual e preferências dos usuários. A Figura 10 apresenta uma visão geral da proposta deste trabalho.



cap_proposta/proposta.pdf

Figura 10 – Visão geral da principal proposta do trabalho.

4.1 Método *VP-Similarity*

Neste trabalho é proposto um método para agrupar os usuários por similaridade de percepção visual. O intuito desse método é manter os usuários que observam imagens de forma semelhante no mesmo *cluster*.

A representação da informação de percepção visual dos usuários é realizada através de um conjunto de fixações. As fixações (Definição 1) de um conjunto de usuários são coletadas através de um rastreador ocular. O rastreador ocular utilizado para coletar os movimentos dos olhos foi o *Tobii X2-60* e sob uma grade de nove pontos para minimizar o ponto da coleta. Ele extraiu o movimento dos olhos, o local e a duração das fixações. Com isso, é possível obter a percepção visual (Definição 2) dos usuários sobre um conjunto de

Tabela 2 – Exemplo da percepção visual sobre três imagens de um conjunto de pinturas. Os asteriscos (*) significa que o usuário não observou e nem avaliou aquela imagem.

	\mathcal{I}_1				\mathcal{I}_2				\mathcal{I}_3			
	q_1	q_2	q_3	q_4	q_1	q_2	q_3	q_4	q_1	q_2	q_3	q_4
u_1	0.50	0.10	0.40	0.00	*	*	*	*	0.15	0.20	0.20	0.45
u_2	0.60	0.20	0.10	0.10	0.10	0.70	0.10	0.05	0.05	0.05	0.10	0.80
u_3	0.40	0.40	0.20	0.00	0.00	0.90	0.10	0.00	*	*	*	*
C_1	0.50	0.23	0.23	0.03	0.05	0.80	0.10	0.05	0.10	0.13	0.15	0.63
u_4	*	*	*	*	0.75	0.08	0.05	0.12	0.85	0.05	0.10	0.00
u_5	0.05	0.25	0.20	0.50	0.70	0.05	0.10	0.15	*	*	*	*
u_6	0.15	0.20	0.25	0.40	0.82	0.02	0.10	0.06	0.70	0.15	0.05	0.10
u_7	*	*	*	*	*	*	*	*	0.60	0.20	0.20	0.00
C_2	0.10	0.23	0.23	0.45	0.76	0.05	0.08	0.11	0.72	0.13	0.12	0.03

imagens. Por meio da percepção visual pode-se determinar a similaridade entre usuários que observaram a mesma imagem, através de um algoritmo de *clusterização*.

Definição 1 (Fixação) Uma fixação f de um usuário u sobre uma imagem I é dada por $f = (p, t)$, onde p é a posição correspondente ao centroíde de um cluster formado por um conjunto de pixels (descrito na Seção 4.4) e t é o tempo de duração da fixação.

Definição 2 (Percepção Visual) Seja o conjunto de fixações $F_{u,I} = \{f_1, \dots, f_n\}$, onde cada $f_i = (p_i, t_i)$ é uma fixação do usuário u na imagem I , para $1 \leq i \leq n$. Considere uma imagem I representada por um conjunto de quadrantes $Q = \{q_1, \dots, q_m\}$. Seja $F_{u,I}^q = \{f_j, \dots, f_k\}$ o conjunto de todas as fixações do tipo $f_l = (p_l, t_l)$, onde p_l está localizada no quadrante q , para $q \in Q$, $F_{u,I}^q \subseteq F_{u,I}$ e $j \leq l \leq k$. A proporção de tempo que o usuário u fixou os olhos em um dado quadrante q é dada por $v = \frac{\sum_{l=j}^k t_l}{\sum_{i=1}^n t_i}$. A percepção visual $P_{u,I}$ de um usuário u sobre a imagem I é dada por tupla com m elementos:

$$P_{u,I} = (v_1, \dots, v_m), \quad (1)$$

onde v_s é a proporção de tempo que o usuário u fixou os olhos no quadrante q_s , para $1 \leq s \leq m$.

A Tabela 2 apresenta um exemplo de como é feita a *clusterização*. Essa Tabela é constituída de 7 usuários ($u_1, u_2, u_3, u_4, u_5, u_6$ e u_7) que são agrupados em dois clusters de percepção visual utilizando a similaridade de percepção visual dos usuários com base em três imagens. Os (*) na tabela indicam que o usuário não viu e nem avaliou aquela imagem. Assim, são obtido dois *clusters* de percepção (Figura 11): no *cluster* 1 estão u_1, u_2 e u_3 enquanto o *cluster* 2 é composto por u_4, u_5, u_6 e u_7 . Além disso, também são exibidos os vetores de consenso de cada cluster que são representados por C^1 e C^2 .

	\mathcal{I}_1				\mathcal{I}_2				\mathcal{I}_3			
	q_1	q_2	q_3	q_4	q_1	q_2	q_3	q_4	q_1	q_2	q_3	q_4
u_1	0.50	0.10	0.40	0.00	*	*	*	*	0.15	0.20	0.20	0.45
u_2	0.60	0.20	0.10	0.10	0.10	0.70	0.10	0.05	0.05	0.05	0.10	0.80
u_3	0.40	0.40	0.20	0.00	0.00	0.90	0.10	0.00	*	*	*	*
C_1	0.50	0.23	0.23	0.03	0.05	0.80	0.10	0.05	0.10	0.13	0.15	0.63
u_4	*	*	*	*	0.75	0.08	0.05	0.12	0.85	0.05	0.10	0.00
u_5	0.05	0.25	0.20	0.50	0.70	0.05	0.10	0.15	*	*	*	*
u_6	0.15	0.20	0.25	0.40	0.82	0.02	0.10	0.06	0.70	0.15	0.05	0.10
u_7	*	*	*	*	*	*	*	*	0.60	0.20	0.20	0.00
C_2	0.10	0.23	0.23	0.45	0.76	0.05	0.08	0.11	0.72	0.13	0.12	0.03

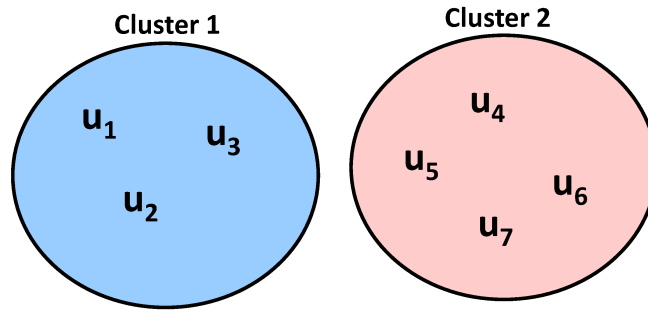


Figura 11 – Exemplo da clusterização por percepção visual. Em destaque de azul estão os usuários do *cluster* 1 e os de vermelho são os usuários do *cluster* 2.

4.2 VP-PrefRec: uma extensão do PrefRec

O VP-PrefRec, como foi dito anteriormente, é a junção do método VP-Similarity com o PrefRec. A escolha do PrefRec se deu pelo fato de ser um *framework* desenvolvido pelo grupo de pesquisa a qual este trabalho está inserido. Por questão de tempo só foi possível fazer a junção do nosso método proposto com este framework, mas futuramente pretende-se usar outros *frameworks* para que assim o nosso método possa ser validado novamente.

As diferenças entre o VP-PrefRec e o PrefRec são: a forma como é construído o modelo de recomendação e como é feita a recomendação. No PrefRec, o processo de construção do modelo de recomendação utiliza avaliações (notas) de usuários e atributos de itens para definir os grupos consensuais e fazer a mineração de preferências em pares. No VP-PrefRec as mesmas atividades são feitas, porém antes da realização dessas atividades é executado um processo de agrupamento baseado na percepção visual, como mostra a Figura 12.

O VP-PrefRec utiliza perfis de usuários com n avaliações de acordo com a seguinte estratégia: o método VP-similarity clusteriza os usuários com base na similaridade de percepção visual. Em seguida são construídos modelos de recomendação para cada *cluster* de percepção visual. Assim, para um novo usuário, é selecionado um *cluster* de percepção visual e, baseado em algumas avaliações o modelo de recomendação, é escolhido e aplicado

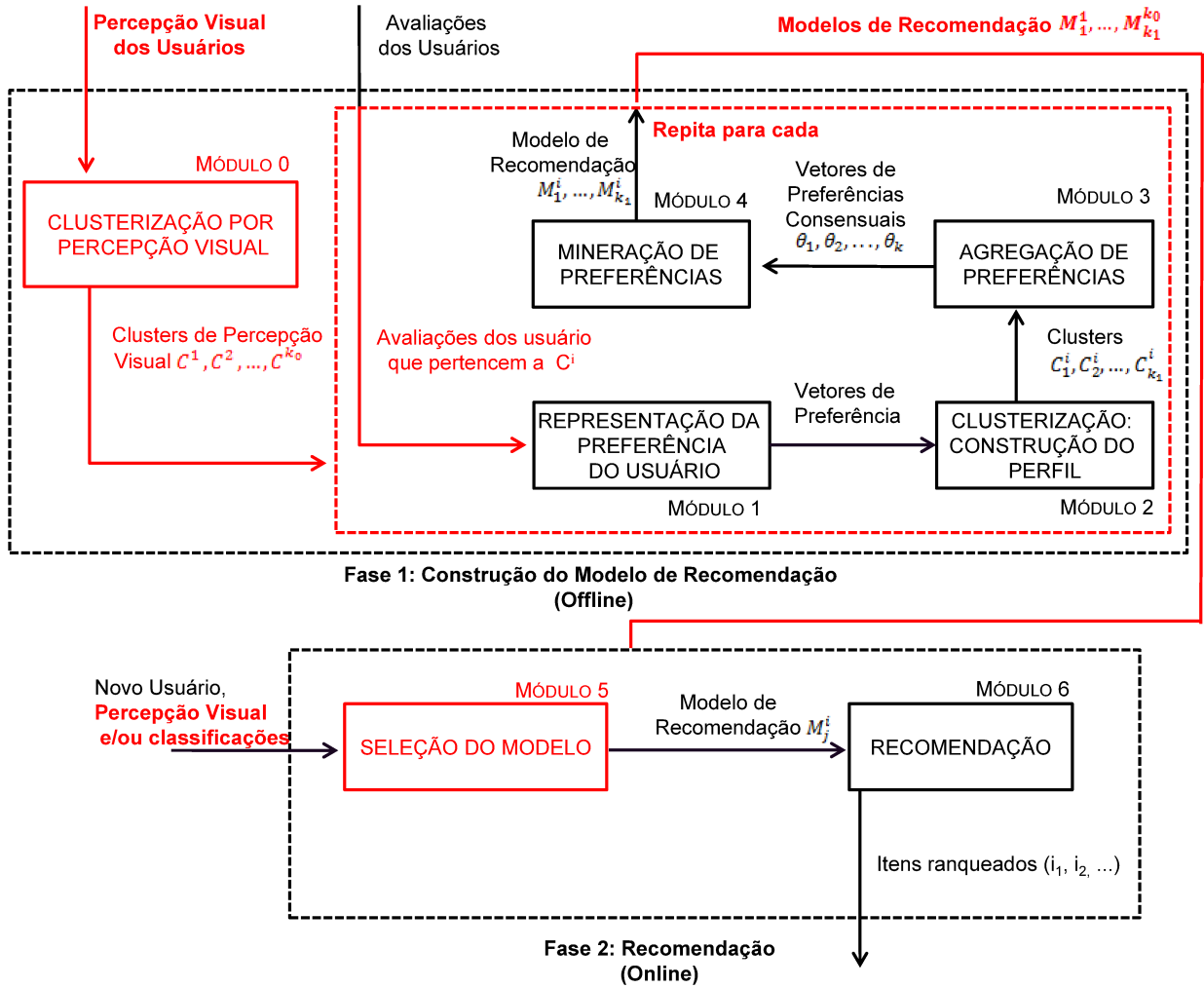


Figura 12 – Arquitetura geral do *VP-PrefRec*. As áreas destacadas de vermelho são as contribuições deste trabalho.

para prever a preferência desse usuário sobre itens ainda não avaliados.

No processo de recomendação é considerado que um novo usuário irá observar e avaliar alguns itens antes da recomendação. A seleção do modelo de recomendação é feita de acordo com similaridade de percepção visual e similaridade de perfil entre um usuário alvo e os *clusters* de preferência consensual. A similaridade de perfis pode ser calculada por qualquer função de similaridade como a distância euclidiana, a similaridade de cosseno, correlação de Pearson, etc. Desse modo, esse processo é executado usando a seguinte abordagem:

1. Por meio um usuário u_t com a informação de percepção visual sobre um conjunto de imagens \mathcal{V} é obtido um vetor de percepção visual α_u correspondente a cada imagem I_u ;

2. É calculado a similaridade entre um usuário α_u e cada vetor consensual de percepção visual δ_i . O usuário α_u é inserido no *cluster* que possui o vetor consensual de percepção visual δ_u mais próximo a ele;
3. Dentro do *cluster* representado por δ_u encontre o modelo de recomendação. Em seguida, diante de um pequeno conjunto de avaliações R_u dada por um usuário u sobre alguns itens I , obtenha o vetor de preferência σ_u correspondente a R_u ;
4. É calculada a similaridade entre σ_u e cada vetor consensual de preferência θ_i , sendo selecionado o vetor consensual θ_u mais próximo de σ_u ;
5. É considerado o modelo de preferência P_u correspondente a θ_u ;
6. P_u é usado para inferir preferências entre pares de itens I que não foram avaliados pelo usuário u no passado. Esses conjuntos de pares de itens (i, j) indicam que o usuário u prefere o item i ao item j , assim, um *ranking* pode ser obtido através da aplicação de um algoritmo de *ranking*.

4.3 Inferindo uma rede de percepção visual

Os sistemas de recomendação tradicionais geralmente utilizam perfis de usuários juntamente com avaliações prévias para predizer itens que irão satisfazer um determinado usuário (WEI; HUANG; FU, 2007; LOPS; GEMMIS; SEMERARO, 2011; BARRILERO et al., 2011). Porém, quando um novo usuário começa a fazer uso do sistema ele não possui nenhum item avaliado. Assim, fica difícil predizer preferências para esse usuário. Este problema é conhecido como *cold start* e é um dos grandes desafios dos sistemas de recomendações (CHOI; JEONG; JEONG, 2010; BERNARDI et al., 2015).

Várias abordagens tentam inferir informações contextuais implícitas de novos usuários com o intuito de contornar o problema do *cold start*. Como exemplo de informação contextual, é possível citar informações sociais (FELÍCIO et al., 2015), o comportamento do usuário baseado no *click* (LIU; DOLAN; PEDERSEN, 2010) e informações baseadas em localização (ZHOU et al., 2012). Notavelmente, com o surgimento das redes sociais, a informação social têm sido amplamente explorada para mitigar o problema do *cold start* na recomendação de itens (LI et al., 2015). Algumas abordagens estão focadas em conteúdos escritos por amigos na rede social (LI et al., 2015). Outros, sobre laços de amizades (LIU; LEE, 2010). No entanto, é importante destacar que nem sempre a informação social é disponível ou é dispendioso para ser acessado (KUROIWA; UDA, 2012).

A percepção visual dos usuário é uma fonte interessante de informação. A ideia de capturar os movimentos dos olhos dos usuários por meio de um rastreador ocular está se tornando cada vez mais tangível com a disponibilidade de acessar essas informações,

por exemplo, por meio da câmera de smartphones (CUONG; HOANG, 2010). E aproveitando o volume de abordagens existentes que tratam o problema do *cold start* através da informação de rede social (CASTILLEJO; ALMEIDA; IPIÑA, 2012; JAMALI; ESTER, 2010), neste trabalho é proposto um método inovador que ao invés de usar uma rede social, faz uso de uma Rede de Percepção Visual inferida a partir das semelhanças em usuários de percepção visual.

Portanto, o *VP-Similarity* foi utilizado para inferir redes de percepção visual dos usuários obtidos a partir de um *cluster*. Usuários que estão no mesmo *cluster* de similaridade de percepção podem ser usados como provedores de recomendação e o grau ligação entre eles pode ser medido através dos *scores* do *VP-similarity*. Com isso, é possível fazer recomendações para usuários sem nenhuma avaliação.

A seleção do Modelo de Recomendação para usuários *cold start* é feita da seguinte forma: primeiramente é determinado o *cluster* de percepção do usuário alvo e então é aplicada uma métrica social que irá selecionar um modelo de recomendação a partir de um cluster com mais usuários que satisfaçam um *threshold* mínimo com grau de ligação entre ele e o usuário alvo.

4.4 Criação da base de dados

A base de dados criada consiste em um conjunto de 607 imagens de obras de artes (pinturas) retiradas do site (<http://pintura.aut.org/>). A Figura 13 apresenta exemplos de imagens que compõem a base de dados. A principal motivação para a criação dessa base de dados foi o fato de não existir até o início dessa pesquisa, uma base de dados que continha tanto dados de preferência como dados de percepção visual dos usuários. Essas pinturas são divididas em 15 classes: animais, arquitetura, arte abstrata, mitologia, natureza morta, natureza morta flores, natureza morta frutas, nudismo, paisagens, paisagens marinhas, criança, mulher, homem, pessoas e religioso. Essas classes foram criadas visualmente baseada nas classificações contidas no site e pode acontecer de existir sobreposição entre as classes, ou seja, uma imagem pode pertencer tanto a classe x como a classe y . A pintura conhecida como *O nascimento de Vênus* é um exemplo de uma imagem que pode pertecer a duas classes que são: nudismo e mitologia. Note que apesar das imagens poderem ter sobreposição das classes, na base de dados não existe nenhuma imagem repetida. Cada imagem possui os seguintes atributos: época, movimento artístico, país obra, pintor e tipo da obra, como mostra a Figura 14.

Foram disponibilizadas 200 imagens aleatórias para cada usuário avaliar com notas de 1 a 5. Essa nota equivale a quanto o usuário gostou da imagem, sendo a nota proporcional ao gosto, ou seja, quanto maior a nota, mais o usuário gostou da imagem. Um rastreador ocular (*eye tracking*) foi utilizado para rastrear os pontos de visão (fixações) do usuário na imagem. Além disso, dados inerentes a cada imagem utilizada durante um



Figura 13 – Algumas imagens da base de dados.

Nome do Atributo	Valores do Atributo
Época	Siglos XVII, Siglos XIX, ... (19 variações)
Movimento Artístico	Renascimento, Barroco, ... (38 variações)
País Obra	Itália, França, ... (13 variações)
Pintor	Nome do pintor da obra (286 variações)
Tipo da Obra	Paisagem Marinha, Natureza morta... (15 variações)

Figura 14 – Atributos das imagens e a quantidade de variações de cada um dos atributos. Essas variações vão influenciar na hora do *VP-PrefRec* construir os modelos de recomendação.

experimento foram armazenados. Com isso foi gerado dois arquivos de saída (.tsv e .xlsx) com os dados brutos de cada experimento. Vale ressaltar, que durante os experimentos tivemos dificuldade de conseguir voluntários para realizar os testes e a maioria dos que se voluntariaram não tiveram paciência para observar todas as imagens com calma e atri-

buir uma nota. Com isso, além da base de dados, os dados de percepção visual também ficaram esparsos.

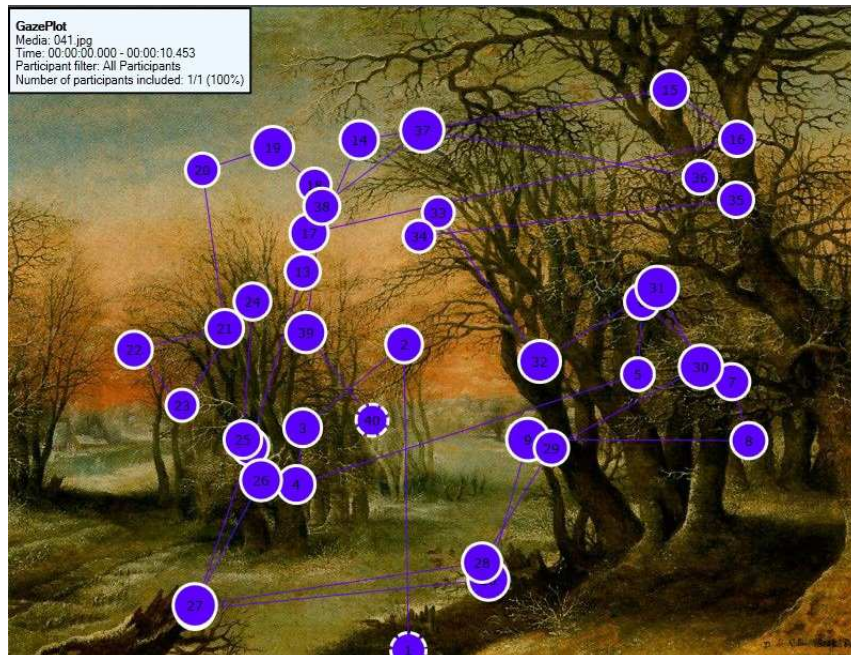


Figura 15 – Atributos de cada imagem.

A Figura 15 ilustra uma das imagens de um experimento realizado utilizando o rastreador ocular. Os círculos roxos (*cluster*) na imagem servem para ilustrar como o usuário observou a imagem durante todo o experimento. De acordo com (STUDIO, 2012), o *cluster* é uma demonstração gráfica das áreas com grande concentração de olhar, aplicado a uma imagem de fundo. Os conjuntos de dados brutos do olhar que sugerem áreas de interesses, podem ser calculados por este tipo de visualização. É utilizado um algoritmo de agrupamento simples para calcular a distância entre dois pontos e deixá-los em um mesmo *cluster* se a distância entre os pontos for inferior a um limite pré-determinado. O algoritmo utilizado por esse rastreador ocular foi proposto por (SANTELLA; DECARLO, 2004) para o estudo dos movimentos dos olhos.

Cada círculo roxo com seu respectivo número mostra a posição e a sequência das fixações em uma imagem estática que foi dada para o usuário observar e avaliar o quanto gostou. O tamanho dos círculos podem variar: quanto maior o círculo roxo, mais tempo o usuário ficou observando aquela região da imagem.

A Figura 16 representa o mapa de calor do teste mostrado na Figura 15. Um mapa de calor utiliza diferentes cores para exibir o número de fixações feitas pelo usuário em determinadas áreas da imagem ou por quanto tempo ele fixou o olhar para determinada região. A cor vermelha mostra o maior número de fixações ou mais tempo. Já a cor verde indica o contrário, menos tempo. Além dessas cores há uma variação no níveis entre essas duas cores.

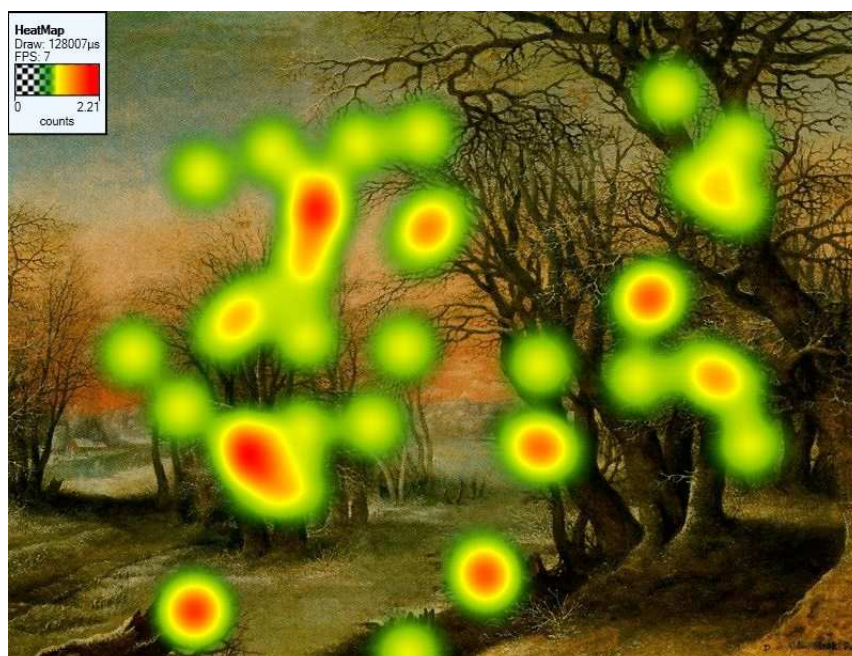


Figura 16 – Mapa de calor do teste realizado na Figura 15.

Resultados Experimentais

Este capítulo tem por objetivo apresentar a abordagem proposta no Capítulo 4 mostrando os resultados alcançados e os experimentos propostos.

5.1 Pré-Processamento da base de dados

A criação da base de dados contou com a participação de 194 usuários com idades entre 18 e 60 anos. Como mencionado na Seção 4.4, cada usuário avaliou um total de 200 imagens aleatórias. Cada uma das 200 imagens foram apresentadas aos usuários para ele observar e avaliar com notas de 1 a 5 de acordo com sua preferência. O tempo médio de duração de cada teste foi de 20 minutos. A base de dados criada possui 38800 avaliações e a esparsidade dos dados é de 67%.

A cada teste realizado com um usuário foi gerado um arquivo de saída contendo os dados do experimento. Esse arquivo possui 56 atributos no total, dos quais foram utilizados apenas 16 atributos. Isso porque cada imagem foi dividida em n -partes e os atributos selecionados contribuíram para o cálculo da porcentagem de tempo que o usuário ficou olhando para cada quadrante da imagem. Assim, os demais atributos foram desconsiderados por não conter informações relevantes para essa análise. A Tabela 3 mostra os 16 atributos selecionados e a descrição de cada um.

Foi feito o pré-processamento desses dados gerados pelo rastreador ocular e com isso, foi criado um arquivo que armazena as notas de preferências dos usuários e os dados de percepção visual.

5.2 Divisão das imagens em q -partes

Cada imagem foi dividida em um *grid* de q partes. Em cada uma dessas partes foi associado a porcentagem de tempo que o usuário ficou observando aquela região da imagem, sendo esse valor foi armazenado em um vetor. Assim, para cada usuário u temos

Tabela 3 – Tabela com os atributos do arquivo de saída

Atributo	Descrição
StudioProjectName	Nome do Projeto.
ParticipantName	Nome do Participante.
Recording Date	Nome do Projeto.
RecordingDuration	Duração total do teste (milissegundos).
MediaName	Nome da mídia (imagem).
KeyPressEvent	Armazena a tecla que o usuário pressionou para avaliar a imagem (NumPad1, NumPad2, NumPad3, NumPad4 ou NumPad5).
FixationIndex	Índice das fixações (auto incremento).
SaccadeIndex	Índice dos movimento dos olhos (auto incremento).
GazeEventType	Indica o evento - fixações, saccade (movimentos dos olhos) ou unclassified (quando não for possível fazer a classificação) - que o equipamento conseguiu capturar.
GazeEventDuration	Duração total de cada evento (milissegundos).
GazePointX (MCSpX)	Média da coordenada horizontal do olho esquerdo e direito na imagem.
GazePointY(MCSpY)	Média da coordenada horizontal do olho esquerdo e direito na imagem.
PupilLeft	Tamanho estimado da pupila esquerda (milímetros).
PupilRight	Tamanho estimado da pupila direita (milímetros).
ValidityLeft	Indica o nível de confiança de que o olho esquerdo foi indentificado corretamente. Os valores variam de 0 (alta confiança) a 4 (olho não encontrado).
ValidityRight	Indica o nível de confiança de que o olho direito foi indentificado corretamente. Os valores variam de 0 (alta confiança) a 4 (olho não encontrado).

associado um conjunto de m vetores de q posições cada um (p_1, \dots, p_q) , sendo um vetor para cada uma das m imagens avaliadas pelos usuários durante a criação da base de dados. Como cada usuário avaliou 200 imagens, temos um total de 200 vetores para cada usuário.

Neste trabalho, foram feitas as divisões das imagens em 4, 9, 16 e 25 partes iguais. O objetivo da variação da quantidade de q é verificar qual o tamanho do vetor que gera melhores *clusters* de percepção. A Figura 17 mostra um exemplo de uma imagem e suas divisões. Já a Figura 18 ilustra como são os vetores criados para cada imagem.

Após a criação dos 200 vetores m de cada usuário, eles são aglutinados de modo a formar um único vetor M que representa um determinado usuário. Uma particularidade desse vetor M é que além dele ser formado por 200 vetores m , ele também é composto por mais 407 vetores de tamanho q , totalizando 607 vetores m . Isso foi necessário para poder medir a similaridade entre dois ou mais usuários, pois as imagens vistas pelos usuário não são a mesma, uma vez que, a seleção da base de teste era feito de forma aleatória (como descrito na Seção 4.4). Por exemplo, os usuários u_1 e u_2 podem ter visto 60 imagens em

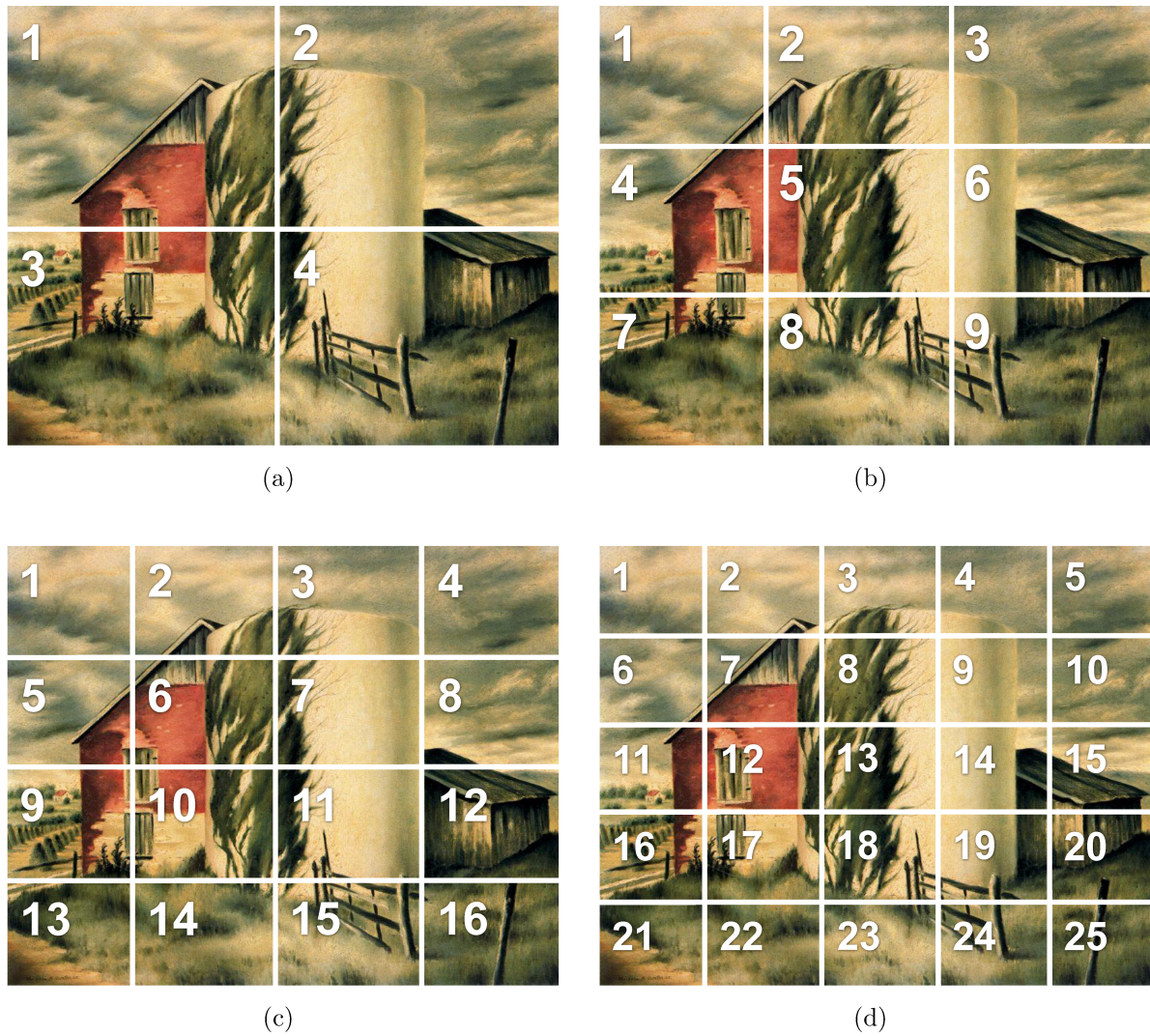


Figura 17 – Exemplos de uma imagem dividida em q partes: (a) $q = 4$; (b) $q = 9$; (c) $q = 16$; (d) $q = 25$;

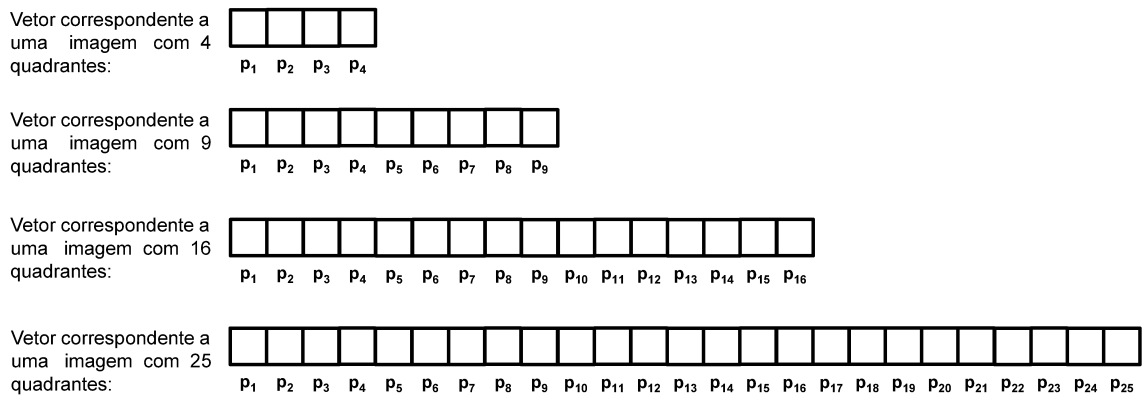


Figura 18 – Vetores criados de cada imagem avaliada por um usuário.

comum, enquanto os usuários u_1 e u_3 viram em comum 90 imagens. Os vetores das 407 imagens que não foram avaliadas pelos usuários foi composto pelo valor 0 (zero) em cada uma das q partes.

5.3 *VP-Similarity*: algoritmo de clusterização dos usuário por percepção visual

O algoritmo de clusterização que foi utilizado para criação dos cluster de percepção visual foi o *k-means*. Nazeer e Sebastian (2009) afirmam que a ideia do algoritmo é classificar um determinado conjunto de dados em um número k de clusters disjuntos, sendo que o valor de k é determinado antecipadamente. O Algoritmo 1 mostra o pseudo código do *k-means*, adaptado de (HAN; KAMBER; PEI, 2011):

Algoritmo 1 *K-means*. O *k-means* é um algoritmo para particionamento, onde cada centróide de um *cluster* é representado pelo valor médio dos objetos no *cluster*.

Entrada: k : número de *clusters*; D: um conjunto de dados contidos em n objetos.

Saída: Um conjunto de k *clusters*.

Método:

1. Selecione aleatoriamente k objetos a partir de D, como centros iniciais;
 2. (Re) Atribuir cada objeto ao *cluster* onde o objeto é mais semelhante, baseado no valor médio dos objetos no *cluster*.
 3. Atualize a média de cada *cluster*, ou seja, calcule o valor médio dos objetos para cada *cluster*.
 4. Repita os itens (b) e (c) até que não haja nenhuma alteração.
-

A escolha aleatória dos centróides iniciais gera diferentes tipos de conjuntos (NAZEER; SEBASTIAN, 2009). Com isso, várias tentativas foram feitas pelos estudiosos para melhorar o desempenho do *k-means*.

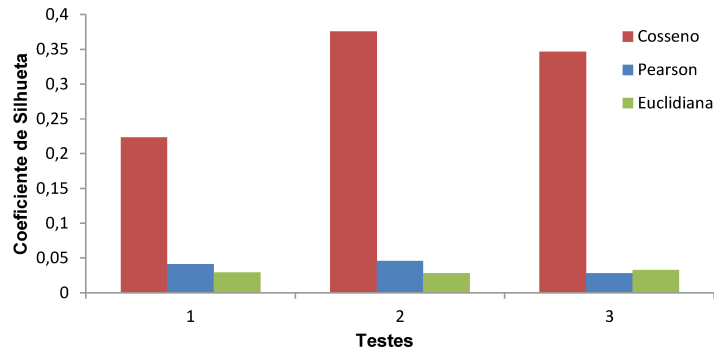
Neste trabalho, foi feita uma adaptação no *k-means* com o intuito de selecionar os centros iniciais mais diferentes para cada grupo. Essa adaptação foi feita usando a estratégia adotada em (PATERLINI; NASCIMENTO; JR, 2011). Assim, a cada execução do *kmeans* os centros iniciais eram o mais distintos um dos outros. Durante todos os experimentos e resultados que serão demonstrados a seguir, o *k-means* foi executado 10 vezes. Dos 10 testes realizados, foram selecionados os 3 melhores resultados.

5.3.1 Seleção da melhor medida de distância

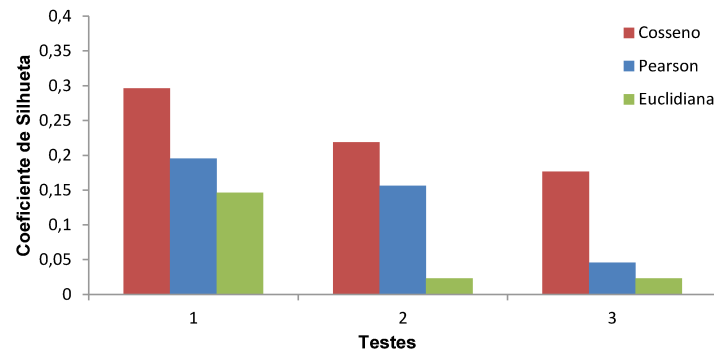
Para calcular a similaridade entre um usuário e um grupo foram realizados testes com diferentes medidas de distâncias. Para tanto, foram utilizadas a distância euclidiana,

Tabela 4 – Exemplo de como é feito o cálculo da distância final entre um usuário e um grupo.

	Situação 1	Situação 2
Imagens	$d(u_1, u_2)$	$d(u_1, u_2)$
1	0,3	<i>Max Value</i>
2	0,25	<i>Max Value</i>
3	<i>Max Value</i>	<i>Max Value</i>
4	0,2	<i>Max Value</i>
5	<i>Max Value</i>	<i>Max Value</i>
Dist. Final	$(0,3 + 0,25 + 0,2) / 3 = 0,25$	<i>Max Value</i>



(a)



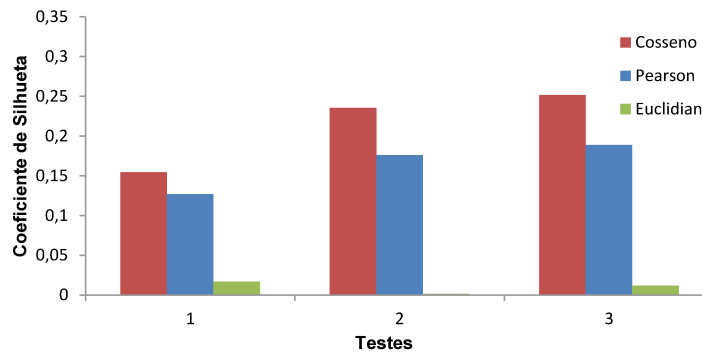
(b)

Figura 19 – Cálculo do Coeficiente de Silhueta para $k = 2$: (a) $q = 4$ e (b) $q = 9$

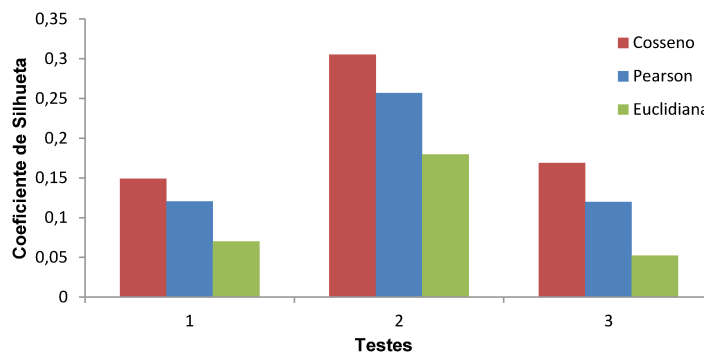
similaridade de cosseno e correlação de pearson. Devido os usuários não terem visto e avaliados as mesmas imagens, o cálculo da distância final é feito como apresentado na Tabela 4. Duas situações podem ocorrer: dois usuários viram pelo menos uma imagem em comum, assim, é calculada a distância entre cada imagem e no final é feito uma

média dessas distâncias, desconsiderando as imagem que ambos não virão (*Max Value*); os usuários não virão nenhuma imagem em comum. Nesse caso, é atribuído um valor muito alto (*Max Value*) na distância final para que esses usuários não fiquem no mesmo grupo.

Foram realizados testes com a quantidade de *clusters* $k = 2$ e $k = 3$. Os dados de entrada para execução do *k-means* foram os arquivos gerados com as divisões das imagens em quadrantes. Assim, os resultados obtidos testando as três métricas são apresentados nas Figuras 19 e 20. Nelas temos dois cenários: o primeiro (Figura 19) foi definido $k = 2$ e testado com os arquivos gerados com as imagens divididas em $q = 4$ e $q = 9$; no segundo cenário (Figura 20), o valor de $k = 3$ e também foram realizados testes com o $q = 4$ e $q = 9$. Para outros valores de k não foi possível gerar *clusters* para nenhuma quantidade q .



(a)



(b)

Figura 20 – Cálculo do Coeficiente de Silhueta para $k = 3$: (a) $q = 4$ e (b) $q = 9$

Para verificar a qualidade dos clusters gerados nestes experimentos, utilizamos o coeficiente de Silhueta. Ele serve para a medir qualidade dos clusters, calculando o quanto um elemento está próximo dos outros elementos do seu grupo e afastado dos elementos

dos outros grupos (HAN; KAMBER; PEI, 2011). Os valores variam de -1 a 1, sendo que melhores resultados são os valores próximo de 1.

Na Figura 19(a) é ilustrado o gráfico de comparação com os três melhores testes das 10 execuções do *k-means*. Nela, é possível perceber que a similaridade de cosseno obteve melhores resultados em todos os testes quando o valor de $q = 4$, em segundo lugar ficou a correlação de Pearson e por último a distância euclidiana. Já na Figura 19(b) quando a imagem é dividida em $q = 9$ a distância de cosseno também apresenta melhores resultados. Nestes testes, o coeficiente de Silhueta obteve resultados mais próximos um dos outros para as três métricas do que se comparado a Figura 19(a), onde a similaridade de cosseno apresentou nitidamente resultados melhores.

Na Figura 20, foi alterado o valor para $k = 3$. Para ambos os valores de q a similaridade de cosseno apresentou melhores resultados. Um detalhe a ser observado é que a distância euclidiana para este tipo de dados não obteve resultados bons em praticamente nenhum teste realizado quando o $q = 4$. No Teste 2, por exemplo, o valor do coeficiente Silhueta obtido com a distância euclidiana foi tão próximo de zero que não dá para ser visto no gráfico, sendo que nesse teste o valor foi de 0,001 %.

Diante dos resultados apresentados, definimos a similaridade de cosseno como sendo a mais adequada para a criação dos *clusters* de percepção visual.

5.3.2 Seleção da melhor configuração de quadrantes

Tabela 5 – Resultados da execução do *k-means* com o $k = 2$

Teste1		
	$q = 4$	$q = 9$
Qtd. de Elementos do Cluster1	169	189
Qtd. de Elementos do Cluster2	25	5
Coeficiente de Silhueta	0,13268	0,16161

Teste2		
	$q = 4$	$q = 9$
Qtd. de Elementos do Cluster1	172	191
Qtd. de Elementos do Cluster2	22	3
Coeficiente de Silhueta	0,12454	0,191725

Teste3		
	$q = 4$	$q = 9$
Qtd. de Elementos do Cluster1	163	192
Qtd. de Elementos do Cluster2	31	2
Coeficiente de Silhueta	0,34628	0,21908

O critério utilizado para verificar se os *clusters* de percepção estão de boa qualidade foi através do cálculo do coeficiente de Silhueta e pela quantidade de elementos em cada

cluster. O segundo critério foi baseado na proposta deste trabalho que consiste em agrupar usuário por percepção visual e depois em cada *cluster* de percepção visual, agrupá-los por preferência. Para isso, é necessário que os clusters gerados com a percepção visual tenha uma quantidade mínima de 10 elementos para que assim possam ser agrupado novamente.

Tabela 6 – Resultados da execução do *k-means* com o $k = 3$

Teste1		
	$q = 4$	$q = 9$
Qtd. de Elementos do Cluster1	8	183
Qtd. de Elementos do Cluster2	8	9
Qtd. de Elementos do Cluster3	178	2
Coeficiente de Silhueta	0,15451	0,12099

Teste2		
	$q = 4$	$q = 9$
Qtd. de Elementos do Cluster1	184	186
Qtd. de Elementos do Cluster2	2	2
Qtd. de Elementos do Cluster3	8	6
Coeficiente de Silhueta	0,23546	0,19652

Teste3		
	$q = 4$	$q = 9$
Qtd. de Elementos do Cluster1	182	2
Qtd. de Elementos do Cluster2	2	2
Qtd. de Elementos do Cluster3	10	190
Coeficiente de Silhueta	0,25147	0,30538

Os resultados apresentados a seguir são para a quantidade de quadrantes $q = 4$ e $q = 9$. Para os valores de $q = 16$ e $q = 25$ o coeficiente de Silhueta foram valores abaixo de 0 ou próximos de 0 e por isso foram desconsiderados.

A Tabela 5 apresenta os resultados da execução do *k-means* com o $k = 2$, com os três melhores resultados das 10 execuções. Nessa tabela é possível perceber que quando a quantidade de quadrantes é $q = 9$, o Teste 1 e Teste 2 apresentam melhores coeficientes de Silhueta do que quando $q = 4$. Porém, a quantidade de elementos que compõe um dos clusters gerados quando $q = 9$ é no máximo de 5 usuários. Assim, esses testes foram desconsiderados por não possuir uma boa quantidade de elementos em cada cluster. Então, quando a quantidade de *clusters* é $k = 2$ foi fixado o valor do $q = 4$.

Na Tabela 6 contém os resultados da execução do *k-means* com o $k = 3$. Independente do valor atribuído a q , um ou mais *clusters* gerados nos testes ficam com poucas quantidades de elementos. Com isso, o segundo critério usado para escolher os melhores *clusters* de percepção não é satisfeito quando o valor de $k = 3$.

5.4 Desempenho do VP-PrefRec

Os resultados do VP-PrefRec foram comparados com o *framework* original PrefRec. A ideia de comparar com este trabalho é verificar se a percepção visual influencia positivamente o *ranking* das recomendações geradas.

Nas duas abordagens, VP-PrefRec e PrefRec, o conjunto de treinamento é usado para construir grupos de usuários similares, definindo um modelo de recomendação para cada grupo. Na fase de teste, para recomendar itens para um usuário, é necessário selecionar o modelo de recomendação do *cluster* mais similar ao usuário, e é nessa etapa que as duas abordagens diferem. Como o sistema PrefRec não utiliza informações de percepção visual no processo de recomendação, a escolha do modelo é realizada através de avaliações prévias de alguns itens, enquanto para o VP-PrefRec é necessário escolher o *cluster* de percepção a qual pertence um novo usuário para depois selecionar o *cluster* de preferência e utilizar o modelo de recomendação gerado por esse *cluster* para poder fazer a recomendação.

O protocolo utilizado no sistema de recomendação foi o *cross-validation* com 5 *folds*. O conjunto de usuários \mathcal{U} foram divididos em 5 partes, assim como o conjunto de itens. A cada iteração uma das partes foi utilizada como usuários de testes (usuários alvo) e as outras quatro partes ficaram sendo o conjunto de usuário de treinamento. Essas cinco iterações são chamadas de iterações dos usuários. Dentro de cada iteração do usuário, para cada usuário alvo u , o conjunto de itens \mathcal{I} foi dividido em 5 partes da seguinte maneira: os itens em \mathcal{I} avaliados por u são distribuídos de forma estratificada em 5 partes de acordo com as avaliações (notas) de u para esses itens. O conjunto de treinamento, com 4 partes de usuários é *clusterizada* com base nas avaliações dos itens e o modelo de recomendação é *contruído*. Na fase de teste, é utilizado as avaliações dos usuários dessas quatro partes selecionar um *cluster* similar por avaliações. Depois de escolher o grupo mais semelhantes, são realizadas as predições de preferências de um item para o outro. Por fim, o restante das notas do usuário alvo são utilizadas para a validação do modelo.

Para avaliar os resultados obtidos com o VP-PrefRec utilizamos o NDCG. O NDCG é uma métrica que serve para medir a qualidade do ranking gerado pelo modelo de recomendação M (JÄRVELIN; KEKÄLÄINEN, 2002). Esse modelo M obtém os pares de itens de testes e prediz qual item é preferido para o usuário alvo. Conforme o resultado da predição é extraído um *ranking* com n itens a serem recomendados para o usuário alvo. Para cada item do *ranking* predito para um usuário é realizado o desconto logaritmo da posição que ele ocupa, obtendo o DCG do ranking predito. Para encontrar o *ranking* ideal o mesmo procedimento é feito, obtendo o DCG*. O cálculo do NDCG é feito através da proporção entre o *ranking* predito (DCG) e o *ranking* ideal (DCG*), produzindo um valor entre 0 e 1, sendo que quanto mais próximo de 1 melhor a qualidade do *ranking* predito. A média aritmética do NDCG dos usuários fornece o valor da qualidade do ranking predito pela abordagem. Note que para cada *cluster* de percepção é obtido um valor de NDCG.

Assim, o valor do NDCG final é feito por meio da média ponderada do NDCG obtido para o *cluster* 1 e *cluster* 2 de percepção visual.

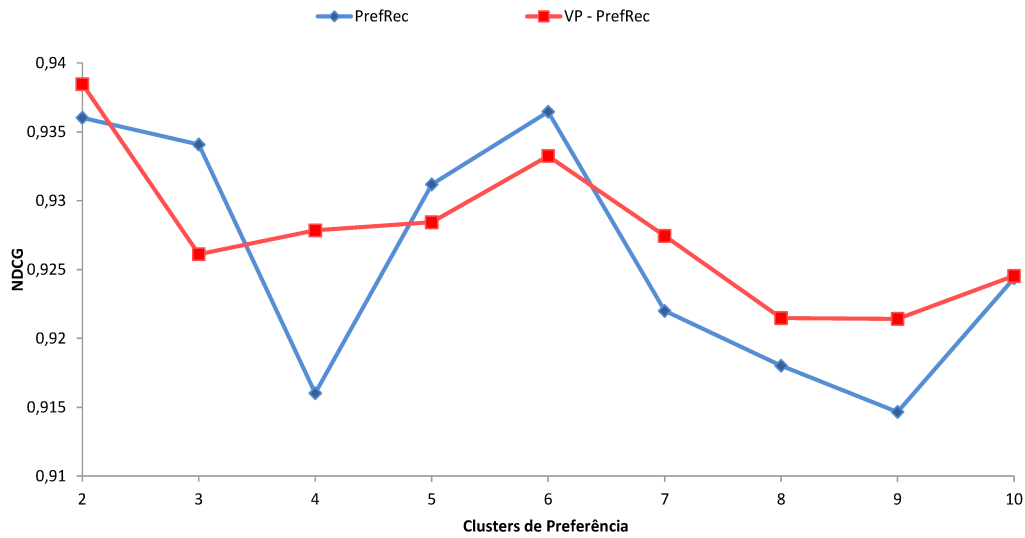


Figura 21 – Resultado do NDCG = 5.

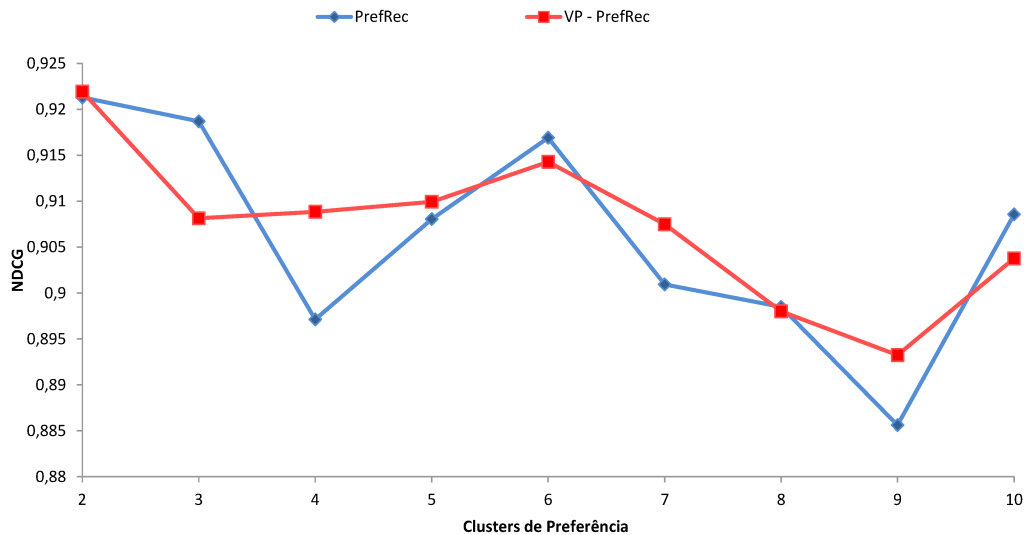


Figura 22 – Resultado do NDCG = 10.

Os testes a seguir foram realizados com os seguintes tamanhos de *NDCG*: 5, 10, 15 e 20. As Figuras 21, 22, 23 e 24 exibem os resultados comparativos com os testes realizado com o *PrefRec* e o *VP-PrefRec*. Para todos as variações do NDCG o *VP-PrefRec* obtém melhores resultados para 6 tamanhos de *clusters* de preferência enquanto o *PrefRec* sobressai apenas para 3 variações de cluster de preferência dos 9 valores diferentes.

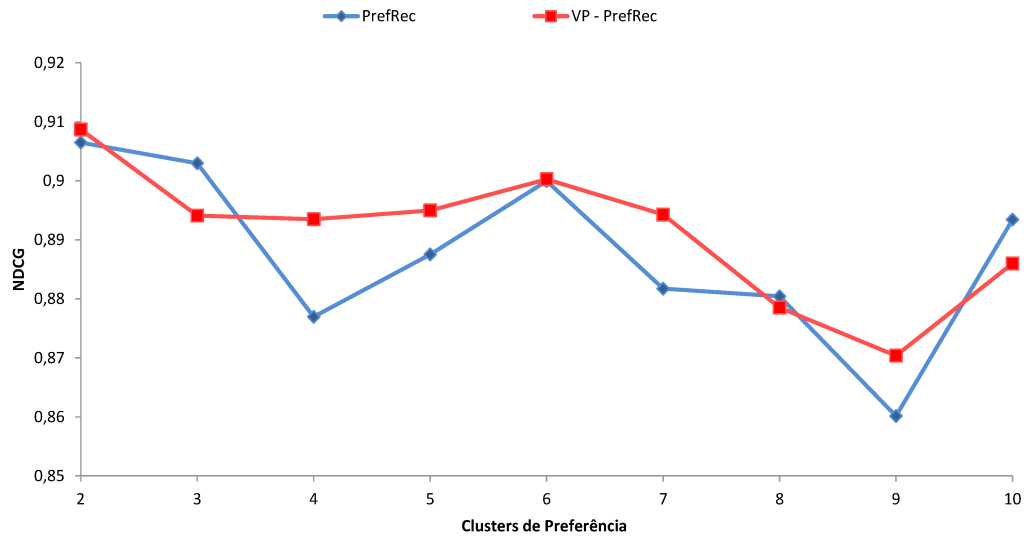


Figura 23 – Resultado do NDCG = 15.

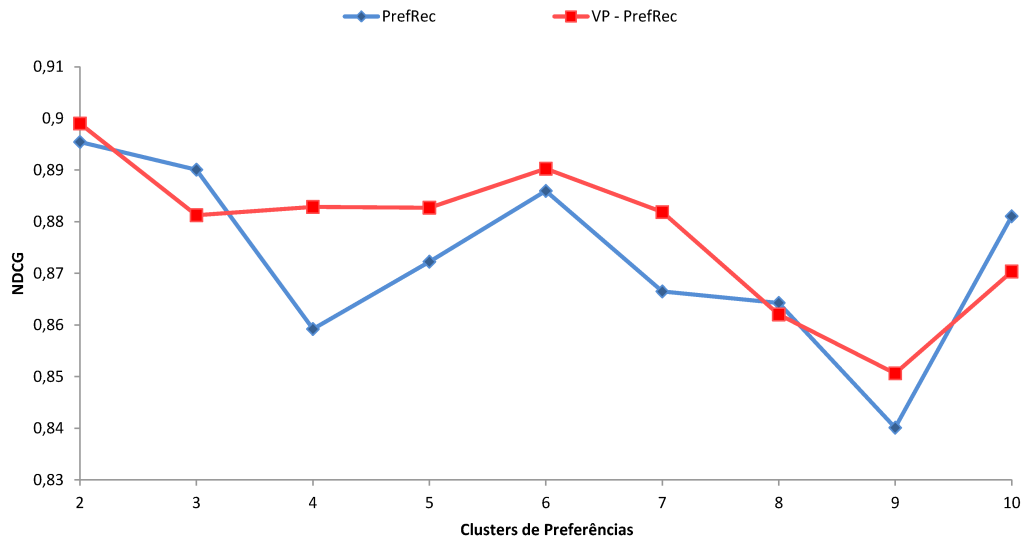


Figura 24 – Resultado do NDCG = 20.

5.5 Influência da inferência da rede de percepção visual

O intuito dos testes realizados utilizando perfis de usuários com nenhuma avaliação foi avaliar se através de uma rede de percepção visual é possível minimizar o problema do *cold start* do usuário. O problema do *cold start* é um desafio enfrentado em praticamente todos os sistemas de recomendação. Ele ocorre quando um novo usuário começa a fazer uso do sistema, pois esse usuário não possui nenhum item avaliado e assim, o sistema

de recomendação tem dificuldades de prever preferências para esse novo usuário.

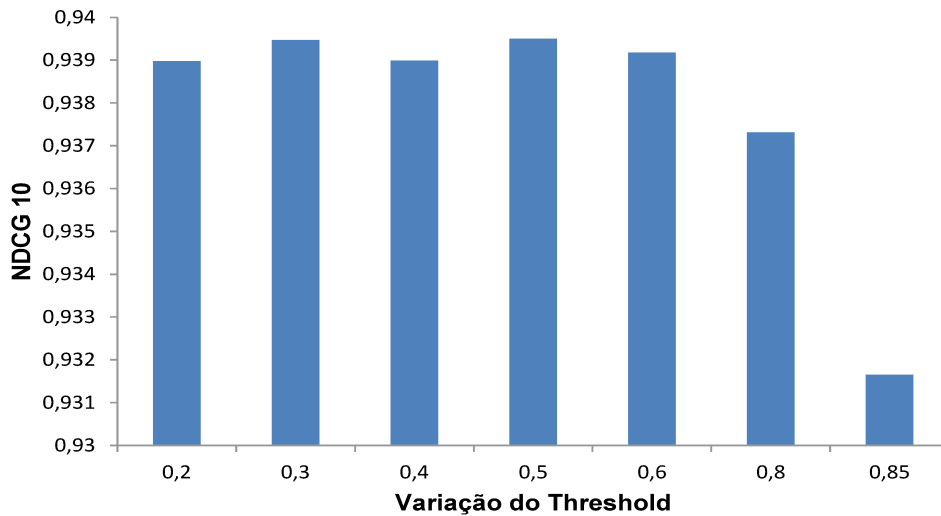


Figura 25 – Variação do *threshold* da métrica *similarity*

A rede de percepção foi aplicada em um sistema de recomendação social chamado Social PrefRec (FELÍCIO et al., 2015). Esse sistema é uma abordagem híbrida e social. Nele, o modelo de recomendação é construído de modo *offline* utilizando usuários em um conjunto de treinamento. É realizada a clusterização dos usuários por avaliações (notas) usando o algoritmo *k-means* e aplicando a mineração de preferências para gerar o modelo de recomendação. A fase de teste incluem as mesmas atividades que o PrefRec.

Os testes foram realizados utilizando três abordagens: duas abordagens utilizando métricas sociais (*similarity* e *friendship*) e a outra utilizando 30 itens avaliados por um novo usuário para poder criar o modelo de recomendação. A métrica *friendship* considera todos os usuário contidos no mesmo *cluster* de percepção como tendo o mesmo peso na hora de criar a rede de percepção. Já com métrica *Similarity* é necessário definir o grau de ligação (similaridade) dos usuário que farão parte de uma mesma rede de percepção visual. Foram testados alguns valores de *threshold* para definir qual o limiar mínimo mais adequado para esta abordagem. A Figura 25 exibe os resultados obtidos com vários *threshold*: 0,2, 0,3, 0,4, 0,5, 0,6, 0,8 e 0,85. Nela é apresentado a média dos valores obtidos para cada valor de *threshold*. O *threshold* = 0,3 e *threshold* = 0,5 foram o que mais se destacaram e com a mesma média para as duas variações. Assim, qualquer um dos dois valores poderiam ser definidos como o *threshold* que seria utilizado nos testes posteriores. Definimos o valor de 0,5 ao invés de 0,3 para termos uma diferença maior entre o *similarity* e o *friendship*, uma vez que no *friendship* todo mundo tem o mesmo peso e no caso do *similarity* só serão considerados os usuários que possuem uma similaridade mínima de 0,5.

Nos testes realizados, variamos o tamanho do *ranking* do NDCG para quatro possíveis: 5, 10, 15 e 20. As Figuras 26, 27, 28 e 29 mostram os resultados comparativos entre três testes realizados com as três abordagens citadas anteriormente. A Figura 26 apresenta os

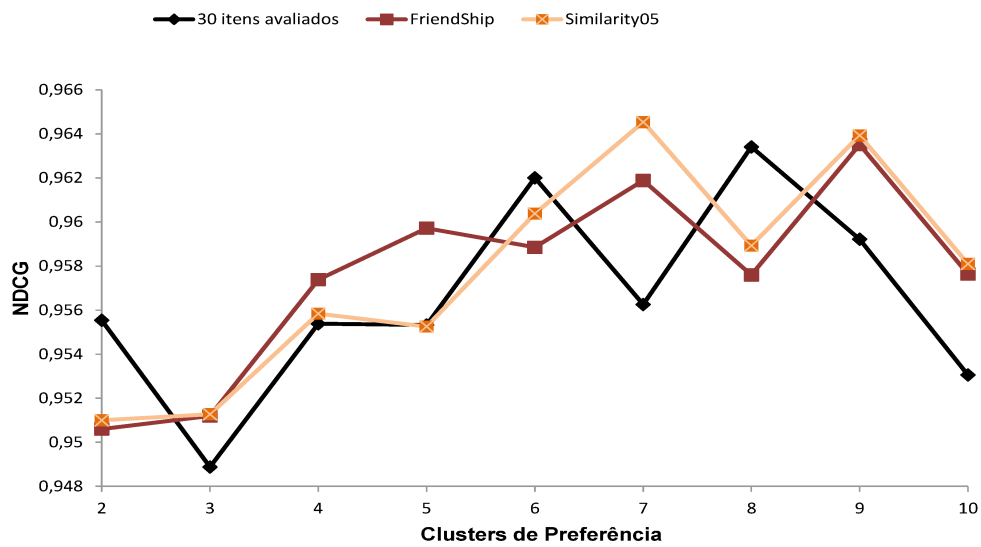


Figura 26 – Resultado do NDCG = 5.

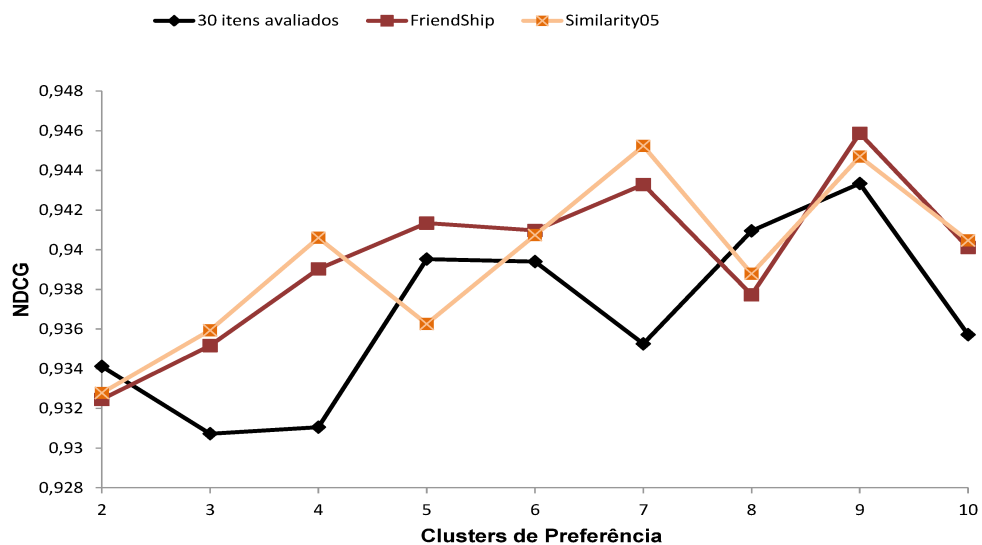


Figura 27 – Resultado do NDCG = 10.

resultados para o $NDCG = 5$. Nesse caso, os resultados obtidos são inconclusivos sobre qual métrica se comportou melhor. Já nos casos onde o *ranking* aumenta para 10 (Figura 27), 15 (Figura 28) ou 20 (Figura 29) itens é possível perceber que as métricas sociais tiveram melhores resultados. As vezes ocorre de em algum dos cluster de preferência o resultado superar a rede de percepção visual, mas ainda assim, os resultados alcançados são significativos, pois consegue recomendar itens para um novo usuário sem necessidade de avaliações prévias.

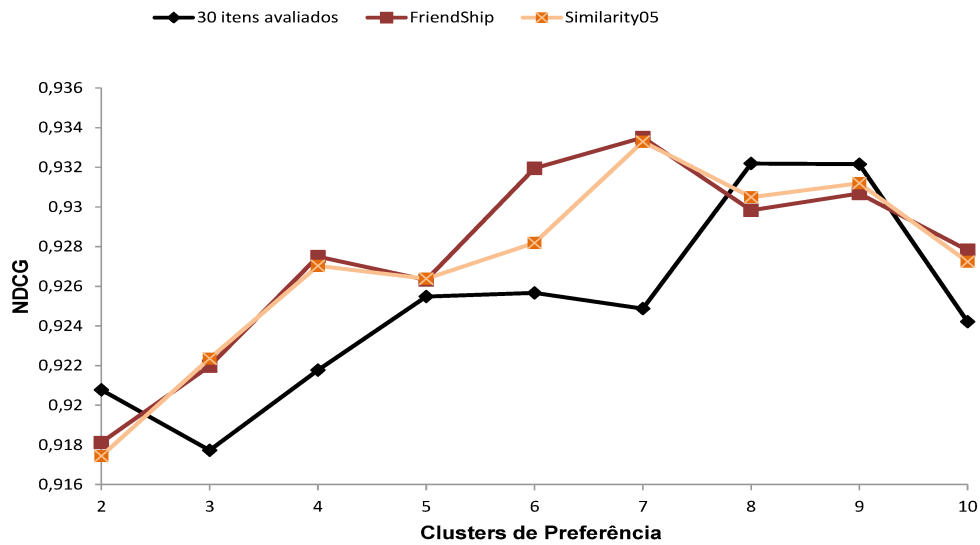


Figura 28 – Resultado do NDCG = 15.

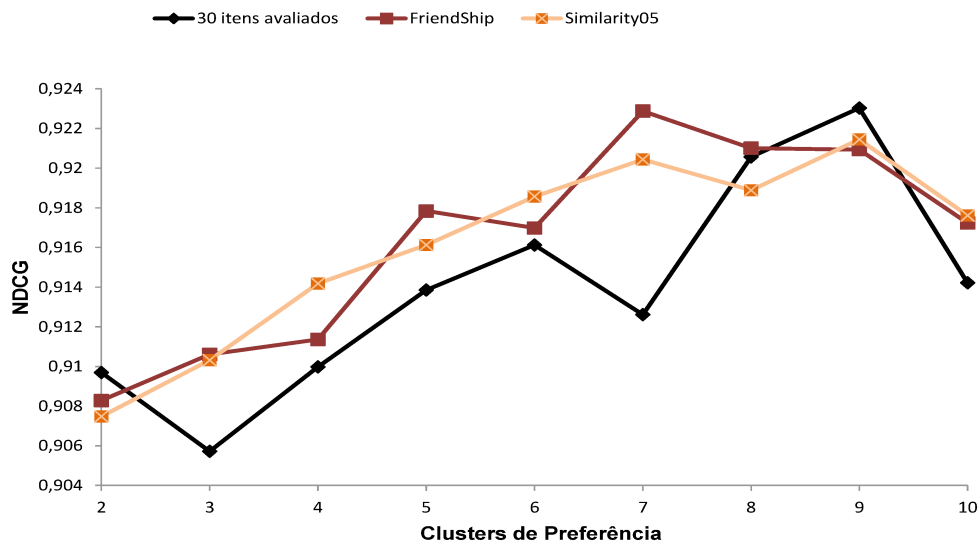


Figura 29 – Resultado do NDCG = 20.

5.6 Influência dos atributos de imagens

Com o intuito de verificar a influência dos atributos de imagens no processo de recomendação foi feita a extração de cores das imagens. Existem diversos modelos de cores que podem ser usado no processamento de imagens. Como descrito na seção 2.3, o modelo RGB é formado pelas cores vermelho, verde e azul. Segundo (), esse modelo é utilizado em aplicação orientadas a hardware, como por exemplo em monitores a cores. Porém, esse modelo não é adequado para descrever percepções humanas, uma vez que as cores não são formadas apenas por essas cores primárias.

O sistema visual humano descreve um objeto de cor pela sua cromaticidade que é definida por matiz e saturação e por seu brilho (CHEN; SHI; XUAN, 2007). Assim, foi utilizado o modelo de cores HSV, que é motivado pelo sistema visual humano. HSV é a sigla para o sistema de cores formado pelas componentes *Hue* (Matiz), *Saturation* (Saturação) e *Value* (Valor). Esse modelo é bem próximo a como os seres humanos conseguem enxergar a cor (GONZALEZ; WOODS, 2010).

Os três componentes desse modelo podem ser definidos como:

- ❑ Matiz: é a cor com saturação e luminosidade máximas. A tonalidade das cores varia entre 0 e 360.
- ❑ Saturação: indica a maior ou menor intensidade.
- ❑ Valor: indica a luminosidade ou brilho de uma cor. Ele indica se é uma imagem clara ou escura.

Neste trabalho, explora-se a matiz (H) e o valor (V) do HSV. A matiz foi extraída com o objetivo de descobrir a cor predominante em cada imagem, sendo elas: cores quentes, cores frias, colorida ou monocromática. Diversos estudos mostram que diferentes cores despertam sentimentos distintos nas pessoas. As cores quentes e frias geralmente são usadas com a intenção de causar essas distintas sensações nas pessoas que as observam. Por exemplo, o vermelho é uma considerado uma cor quente, e geralmente é associada a sentimentos como a paixão ou a raiva. Enquanto o azul é uma cor fria, e é associado a serenidade, bondade. A Figura 30 ilustra um círculo cromático que mostra quais cores são consideradas cores quentes e cores frias.

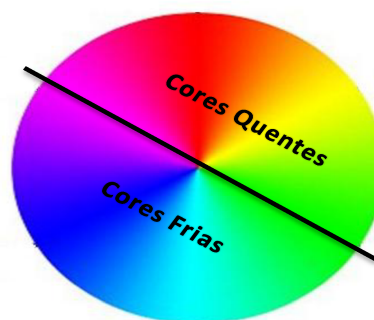


Figura 30 – Círculo cromático representando as cores frias e quentes.

Como no componente matiz do HSV a tonalidade das cores variam de 0° a 360° (Figura 31), o atributo matiz das imagens foram categorizadas por meio da proporção de pixels que estão contidos nos intervalos definidos para cada atributo. Com isso é possível determinar qual a predominância de cores em uma imagem. Um limiar estabelece qual é a proporção mínima necessária. A seguir os intervalos estabelecidos:

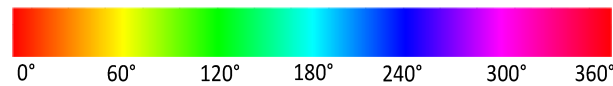


Figura 31 – Barra de cores representadas em graus.

- ❑ Cores quentes: pertencem ao intervalo de 0° a 120° (inclusive) e de 300° a 360° (inclusive);
- ❑ Cores frias: pertecem ao intervalo de 120° e 300° (inclusive);

Neste trabalho, uma imagem que possua pelo menos 65% dos pixels de cores quentes é definida como uma imagem de cores quentes. Se a imagem possui pelo menos 65% dos pixels de cores frias é estabelecido que aquela imagem é de cores frias. Considera-se uma imagem como colorida aquelas que possuem uma proporção de pixels balanceados entre cores quentes e cores frias. Por fim, as imagens monocromáticas são aquelas onde não é possível definir o valor de H (matiz). Isso acontece quando o V (valor) for igual a zero. Nesse caso não há matiz no eixo de intensidade. Então, caso a proporção dos pixels de uma imagem em que o H não pode ser definido seja maior do 50%, considera-se essa imagem monocromática.

O valor (V) foi extraído com a finalidade de descobrir se cada imagem era clara, escura ou possui um bom contraste. Assim, foi gerado o histograma de cores de cada imagem. O brilho de uma imagem pode variar de 0 a 255 para cada cor com uma profundidade de 8-bits. Em um histograma se a proporção do brilho nos pixels for de 0 a 128, essa imagem é considerada como tendo a predominância de cor escura. Caso a proporção do brilho for de 129 a 255 é estabelecido que essa imagem é clara. Um limiar estabelece qual é a proporção mínima necessária. Nesta pesquisa, foi estabelecido um limiar de 65%. Assim, considera-se uma imagem com bom contraste aquela que possui um balanceamento entre as cores clara e escura.

Portanto, através das características extraídas da imagem, juntamente com a imagem e os dados convencionais, foram compostos os dados do objeto que foram utilizados para testar o sistema de recomendação, como mostra a Figura 32.

O resultado obtido usando as características de imagem é apresentado no gráfico da Figura 33. Nessa figura é feita a comparação da acurácia do método proposto usando os atributos de imagem e sem o uso desses atributos. Como é possível observar, os resultados obtidos foram semelhantes aos já existentes. Essa proximidade nos resultados ocorre possivelmente porque foram extraídas características relacionadas a um único componente da imagem que é a cor. No entanto, existem outras formas de extrair características das imagens. Dentro deste contexto, uma proposta para trabalhos futuros é explorar outras

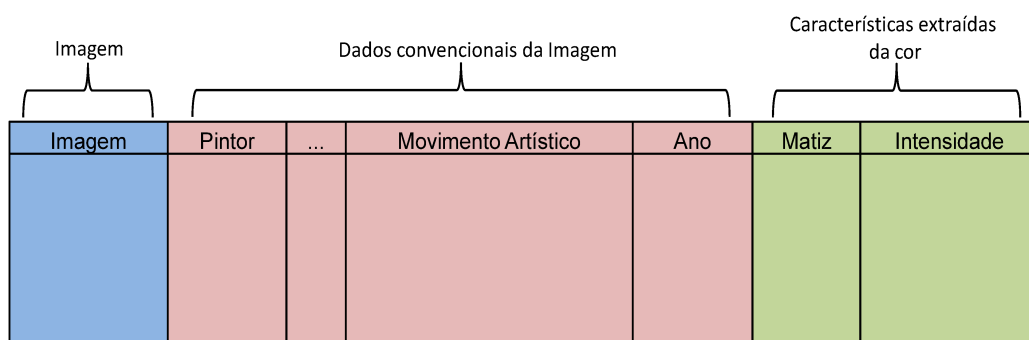


Figura 32 – Dados dos itens.

maneiras de extrair características das imagens, obter e analisar novos resultados a fim de comparar os diversos métodos para a mesma abordagem.

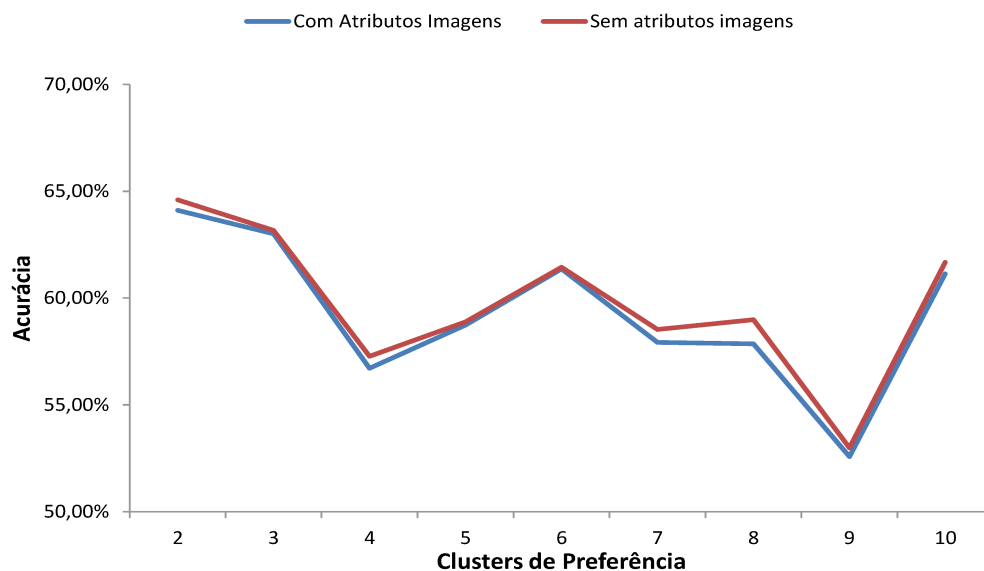


Figura 33 – Comparação Acurácia com o uso de atributos imagens e sem o uso desses atributos.

Conclusão

Neste trabalho foi realizado um estudo sobre a percepção humana e mineração de preferências contextuais na recomendação de imagens para usuários. A pesquisa teve como objetivo mostrar que é possível melhorar os sistemas de recomendação de imagens agregando informações de percepção visual dos usuários. Para isso, foi proposto um método chamado *VP-Similarity* que faz o agrupamento dos usuários por similaridade de percepção visual. A ideia geral desse método é que usuários que observam imagens de forma semelhante estejam no mesmo grupo. Anteriormente, foi construída uma base de dados de percepção visual e de preferências de usuários no domínio de pinturas. Essa base conta com 607 imagens de pinturas divididas em 15 classes: animais, arquitetura, arte abstrata, mitologia, natureza morta, natureza morta flores, natureza morta frutas, nudismo, paisagens, paisagens marinhas, criança, mulher, homem, pessoas e religioso. A criação dessa base teve a participação de 194 voluntário que avaliaram 200 imagens aleatórias com notas de 1 a 5, sendo que a nota é proporcional a quanto ele gostou da imagem. A medida que os usuários avaliavam as imagens, um rastreador ocular (*eye tracking*) capturava os pontos de fixações dos usuários na imagem. Assim, os dados de percepção visual e preferência da base criada foram pré-processados e serviram como dado de entrada para o método *VP-Similarity*.

Foi proposto também a extensão do sistema *PrefRec*, gerando assim, o *VP-PrefRec*. Ele é constituído pela junção do método *VP-Similarity* com o *PrefRec*. Com isso, foi obtido um sistema de recomendação de imagens baseado em dados de percepção visual e preferências dos usuários.

Com o objetivo de minimizar o problema do *cold start* existente em muitos sistemas de recomendação, foi proposto um método que ao invés de usar uma rede social, faz uso de uma rede de percepção visual inferida a partir das semelhanças em usuários de percepção visual. O intuito é que usuários que estão no mesmo *cluster* de similaridade de percepção podem ser usados como provedores de recomendação e o grau ligação entre eles pode ser medido através dos *scores* do *VP-similarity*.

Como resultado, o trabalho apresentou a melhoria do *ranking* das recomendações

geradas. Além disso, os resultados mostram que utilizando a informação de percepção visual para inferir uma rede de percepção é interessante, pois com a rede é possível fazer a recomendação para novos usuários conseguindo minimizar o problema do *cold start*.

6.1 Principais Contribuições

Este trabalho apresenta as seguintes contribuições através das quais os objetivos foram atingidos e as hipóteses validadas:

- ❑ Criação do método ***VP-Similarity*** para representação e cálculo de similaridade entre dois usuários em relação às suas percepções visuais;
- ❑ Extensão do sistema *PrefRec*, gerando assim, o ***VP-PrefRec*** - um sistema de recomendação de imagens baseado em dados de percepção visual e preferências dos usuários;
- ❑ Criação da base de dados ***UFU-Paintings***¹. A base contém dados de preferência e de percepção visual dos usuários no domínio de pinturas;
- ❑ Aplicação do ***VP-PrefRec*** sobre um sistema de recomendação social, através da inferência de uma rede de percepção visual, tratando o problema do *cold-start* de usuário;
- ❑ Uma bateria de testes que comparam o desempenho do sistema ***VP-PrefRec*** com o sistema existente o *PrefRec*;
- ❑ Melhoria do *ranking* das recomendações geradas (hipóteses H1 e H2).

6.2 Trabalhos Futuros

Por se tratar de uma nova área de investigação, este trabalho abre muitos caminhos para trabalhos futuros. A seguir são listados alguns possíveis trabalhos futuros.

- ❑ Explorar um método alternativo ao *VP-Similarity* baseado não em *clusterização*, mas em técnicas de aprendizado supervisionado sobre a percepção visual;
- ❑ Estender a recomendação por percepção visual para além de imagens, como por exemplo, páginas web ou vídeos;
- ❑ Construir uma nova base de dados para que assim os métodos propostos possam ser avaliados em outro domínio, como por exemplo, o domínio de roupas;

¹ Disponível em: www.lsi.facom.ufu.br/datasets

- ❑ Combinar a percepção visual com outros tipos de dados, sem ser preferências e itens. Por exemplo: combinar percepção visual com dados de contexto ou com dados de redes sociais;
- ❑ Explorar a dinâmica da percepção visual a fim de verificar como é a evolução da percepção visual do usuário ao longo do tempo. A ideia é tentar descobrir se um usuário percebe a mesma imagem, duas ou mais vezes, de maneira similar.
- ❑ Explorar mais formas de extração de características de imagens. Por exemplo, extrair característica dos pontos de fixações dos usuários ao invés da imagem toda.

6.3 Contribuições em Produção Bibliográfica

- ❑ Artigo intitulado “Visual Perception Similarities to Improve the Quality of User Cold Start Recommendations”, submetido e aceito na 29th Canadian Conference on Artificial Intelligence.

Referências

- ADOMAVICIUS, G.; TUZHILIN, A. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. **Knowledge and Data Engineering, IEEE Transactions on**, IEEE, v. 17, n. 6, p. 734–749, 2005.
- _____. Context-aware recommender systems. In: **Recommender systems handbook**. [S.l.]: Springer, 2011. p. 217–253.
- AHMAD, S. **VISIT: An efficient computational model of human visual attention**. Tese (Doutorado) — University of Illinois at Urbana-Champaign, 1991.
- AMO, S. de; OLIVEIRA, C. G. Towards a tunable framework for recommendation systems based on pairwise preference mining algorithms. In: **Advances in Artificial Intelligence**. [S.l.]: Springer, 2014. p. 282–288.
- BALABANOVIĆ, M.; SHOHAM, Y. Fab: content-based, collaborative recommendation. **Communications of the ACM**, ACM, v. 40, n. 3, p. 66–72, 1997.
- BARNES, S.; HUNT, B. **E-commerce and v-business**. [S.l.]: Routledge, 2013.
- BARRILERO, M. et al. In-network content based image recommendation system for content-aware networks. In: **Computer Communications Workshops (INFOCOM WKSHPS), 2011 IEEE Conference on**. [S.l.: s.n.], 2011. p. 115–120.
- BERNARDI, L. et al. The continuous cold start problem in e-commerce recommender systems. **arXiv preprint arXiv:1508.01177**, 2015.
- BERRY, S. et al. Netflix recommendations for groups. **Proceedings of the American Society for Information Science and Technology**, Wiley Online Library, v. 47, n. 1, p. 1–3, 2010.
- BHATT, B.; PATEL, P. J.; GAUDANI, H. A review paper on machine learning based recommendation system. In: **IJEDR. International Journal of Engineering Development and Research**. [S.l.], 2014. v. 2, n. 4 (Dec 2014).
- BRODLEY, C. et al. Content-based retrieval from medical image databases: A synergy of human interaction, machine learning and computer vision. In: **AAAI/IAAI**. [S.l.: s.n.], 1999. p. 760–767.
- BURKE, R. Hybrid recommender systems: Survey and experiments. **User modeling and user-adapted interaction**, Springer, v. 12, n. 4, p. 331–370, 2002.

- _____. Hybrid systems for personalized recommendations. In: **Intelligent Techniques for Web Personalization**. [S.l.]: Springer, 2005. p. 133–152.
- _____. Hybrid web recommender systems. In: **The adaptive web**. [S.l.]: Springer, 2007. p. 377–408.
- CANDILLIER, L.; MEYER, F.; BOULLÉ, M. Comparing state-of-the-art collaborative filtering systems. In: **Machine Learning and Data Mining in Pattern Recognition**. [S.l.]: Springer, 2007. p. 548–562.
- CASTILLEJO, E.; ALMEIDA, A.; IPIÑA, D. López-de. Social network analysis applied to recommendation systems: Alleviating the cold-user problem. In: SPRINGER. **UCAmI**. [S.l.], 2012. p. 306–313.
- CHEN, T.; HE, L. Collaborative filtering based on demographic attribute vector. In: IEEE. **Future Computer and Communication, 2009. FCC'09. International Conference on**. [S.l.], 2009. p. 225–229.
- CHEN, W.; SHI, Y. Q.; XUAN, G. Identifying computer graphics using hsv color model and statistical moments of characteristic functions. In: IEEE. **Multimedia and Expo, 2007 IEEE International Conference on**. [S.l.], 2007. p. 1123–1126.
- CHOI, S. H.; JEONG, Y.-S.; JEONG, M. K. A hybrid recommendation method with reduced data for large-scale application. **Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on**, IEEE, v. 40, n. 5, p. 557–566, 2010.
- CUONG, N. H.; HOANG, H. T. Eye-gaze detection with a single webcam based on geometry features extraction. In: IEEE. **Control Automation Robotics & Vision (ICARCV), 2010 11th International Conference on**. [S.l.], 2010. p. 2507–2512.
- DAVIDSON, J. et al. The youtube video recommendation system. In: ACM. **Proceedings of the fourth ACM conference on Recommender systems**. [S.l.], 2010. p. 293–296.
- FELÍCIO, C. Z. et al. Social prefrec framework:leveraging recommender systems based on social information. In: **Proceedings of the 3rd Symposium on Knowledge Discovery, Mining and Learning**. [S.l.: s.n.], 2015. p. 66–73.
- GONZALEZ, R. C.; WOODS, R. E. **Pocessamento digital de imagens**. 3. ed. [S.l.]: Pearson Prentice Hall, 2010.
- HAN, J.; KAMBER, M.; PEI, J. **Data mining: concepts and techniques: concepts and techniques**. [S.l.]: Elsevier, 2011.
- HE, M.; CHI, N. Research on improvement of multi-agent-based e-commerce recommendation system. In: IEEE. **Conference Anthology, IEEE**. [S.l.], 2013. p. 1–4.
- HERLOCKER, J. L. **Understanding and improving automated collaborative filtering systems**. Tese (Doutorado) — University of Minnesota, 2000.
- HERLOCKER, J. L. et al. Evaluating collaborative filtering recommender systems. **ACM Transactions on Information Systems (TOIS)**, ACM, v. 22, n. 1, p. 5–53, 2004.

- HU, Y.; KOREN, Y.; VOLINSKY, C. Collaborative filtering for implicit feedback datasets. In: IEEE. **Data Mining, 2008. ICDM'08. Eighth IEEE International Conference on**. [S.l.], 2008. p. 263–272.
- HUANG, Z.; CHEN, H.; ZENG, D. Applying associative retrieval techniques to alleviate the sparsity problem in collaborative filtering. **ACM Transactions on Information Systems (TOIS)**, ACM, v. 22, n. 1, p. 116–142, 2004.
- HUANG, Z.; CHUNG, W.; CHEN, H. A graph model for e-commerce recommender systems. **Journal of the American Society for information science and technology**, Wiley Online Library, v. 55, n. 3, p. 259–274, 2004.
- ITTI, L.; KOCH, C. Computational modelling of visual attention. **Nature reviews neuroscience**, Nature Publishing Group, v. 2, n. 3, p. 194–203, 2001.
- JAMALI, M.; ESTER, M. A matrix factorization technique with trust propagation for recommendation in social networks. In: ACM. **Proceedings of the fourth ACM conference on Recommender systems**. [S.l.], 2010. p. 135–142.
- JANNACH, D. et al. **Recommender systems: an introduction**. [S.l.]: Cambridge University Press, 2010.
- JÄRVELIN, K.; KEKÄLÄINEN, J. Cumulated gain-based evaluation of ir techniques. **ACM Transactions on Information Systems (TOIS)**, ACM, v. 20, n. 4, p. 422–446, 2002.
- JAWAHEER, G.; SZOMSZOR, M.; KOSTKOVA, P. Comparison of implicit and explicit feedback from an online music recommendation service. In: ACM. **proceedings of the 1st international workshop on information heterogeneity and fusion in recommender systems**. [S.l.], 2010. p. 47–51.
- KIESLER, S. **Culture of the Internet**. [S.l.]: Psychology Press, 2014.
- KOCH, C.; ULLMAN, S. Shifts in selective visual attention: towards the underlying neural circuitry. In: **Matters of Intelligence**. [S.l.]: Springer, 1987. p. 115–141.
- KUO, F.-F. et al. Emotion-based music recommendation by association discovery from film music. In: ACM. **Proceedings of the 13th annual ACM international conference on Multimedia**. [S.l.], 2005. p. 507–510.
- KUROIWA, K.; UDA, R. A low cost privacy protection method for sns by using bloom filter. In: ACM. **Proceedings of the 6th International Conference on Ubiquitous Information Management and Communication**. [S.l.], 2012. p. 7.
- LAM, X. N. et al. Addressing cold-start problem in recommendation systems. In: ACM. **Proceedings of the 2nd international conference on Ubiquitous information management and communication**. [S.l.], 2008. p. 208–211.
- LEE, T. Q.; PARK, Y.; PARK, Y.-T. A time-based approach to effective recommender systems using implicit feedback. **Expert systems with applications**, Elsevier, v. 34, n. 4, p. 3055–3062, 2008.

- LEMOS, F. D. et al. Towards a context-aware photo recommender system. In: **Proceedings of the Context-Aware Recommender System Workshop at ACM Recommender Systems**. [S.l.: s.n.], 2012.
- LI, C. et al. Exploring social network information for solving cold start in product recommendation. In: WANG, J. et al. (Ed.). **Web Information Systems Engineering – WISE 2015**. [S.l.]: Springer International Publishing, 2015, (Lecture Notes in Computer Science, v. 9419). p. 276–283.
- LINDEN, G.; SMITH, B.; YORK, J. Amazon. com recommendations: Item-to-item collaborative filtering. **Internet Computing, IEEE**, IEEE, v. 7, n. 1, p. 76–80, 2003.
- LIU, F.; LEE, H. J. Use of social network information to enhance collaborative filtering performance. **Expert systems with applications**, Elsevier, v. 37, n. 7, p. 4772–4778, 2010.
- LIU, J.; DOLAN, P.; PEDERSEN, E. R. Personalized news recommendation based on click behavior. In: ACM. **Proceedings of the 15th international conference on Intelligent user interfaces**. [S.l.], 2010. p. 31–40.
- LOPS, P.; GEMMIS, M. D.; SEMERARO, G. Content-based recommender systems: State of the art and trends. In: **Recommender systems handbook**. [S.l.]: Springer, 2011. p. 73–105.
- NAZEER, K. A.; SEBASTIAN, M. Improving the accuracy and efficiency of the k-means clustering algorithm. In: **Proceedings of the World Congress on Engineering**. [S.l.: s.n.], 2009. v. 1, p. 1–3.
- OLIVEIRA, C. G.; AMO, S. de. Xprefrec: minimizando o problema de cold-start de item com mineração de preferências. **Journal of Information and Data Management**, 2013.
- PATERLINI, A. A.; NASCIMENTO, M. A.; JR, C. T. Using pivots to speed-up k-medoids clustering. **Journal of Information and Data Management**, v. 2, n. 2, p. 221, 2011.
- POST, F. H. et al. The state of the art in flow visualisation: Feature extraction and tracking. In: WILEY ONLINE LIBRARY. **Computer Graphics Forum**. [S.l.], 2003. v. 22, n. 4, p. 775–792.
- QIU, T. et al. An item-oriented recommendation algorithm on cold-start problem. **EPL (Europhysics Letters)**, IOP Publishing, v. 95, n. 5, p. 58003, 2011.
- REED, T. R.; DUBUF, J. H. A review of recent texture segmentation and feature extraction techniques. **CVGIP: Image understanding**, Elsevier, v. 57, n. 3, p. 359–372, 1993.
- RESNICK, P. et al. Grouplens: an open architecture for collaborative filtering of netnews. In: ACM. **Proceedings of the 1994 ACM conference on Computer supported cooperative work**. [S.l.], 1994. p. 175–186.
- SANCHEZ, F. et al. Social and content hybrid image recommender system for mobile social networks. **Mobile Networks and Applications**, Springer, v. 17, n. 6, p. 782–795, 2012.

- SANDON, P. A. Simulating visual attention. **Journal of Cognitive Neuroscience**, MIT Press, v. 2, n. 3, p. 213–231, 1990.
- SANTELLA, A.; DECARLO, D. Robust clustering of eye movement recordings for quantification of visual interest. In: ACM. **Proceedings of the 2004 symposium on Eye tracking research & applications**. [S.l.], 2004. p. 27–34.
- SARWAR, B. et al. Item-based collaborative filtering recommendation algorithms. In: ACM. **Proceedings of the 10th international conference on World Wide Web**. [S.l.], 2001. p. 285–295.
- SCHEIN, A. I. et al. Methods and metrics for cold-start recommendations. In: ACM. **Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval**. [S.l.], 2002. p. 253–260.
- SHANI, G.; GUNAWARDANA, A. Evaluating recommendation systems. In: **Recommender systems handbook**. [S.l.]: Springer, 2011. p. 257–297.
- STUDIO, T. 3.2 user manual. **Tobii Technology AB**, 2012.
- SUGANO, Y. et al. Image preference estimation with a data-driven approach: A comparative study between gaze and image features. **Journal of Eye Movement Research**, v. 7, n. 3, p. 5, 2014.
- SWAIN, M. J.; BALLARD, D. H. Color indexing. **International journal of computer vision**, Springer, v. 7, n. 1, p. 11–32, 1991.
- TORTORA, G. J.; DERRICKSON, B. **Corpo humano: fundamentos de anatomia e fisiologia**. [S.l.]: Artmed Editora, 2012.
- TREISMAN, A. M.; GELADE, G. A feature-integration theory of attention. **Cognitive psychology**, Elsevier, v. 12, n. 1, p. 97–136, 1980.
- TSOTSOS, J. K. Analyzing vision at the complexity level. **Behavioral and brain sciences**, Cambridge Univ Press, v. 13, n. 03, p. 423–445, 1990.
- UMEMOTO, K. et al. Search intent estimation from user’s eye movements for supporting information seeking. In: ACM. **Proceedings of the International Working Conference on Advanced Visual Interfaces**. [S.l.], 2012. p. 349–356.
- WAN, S.; JIN, P.; YUE, L. An approach for image retrieval based on visual saliency. In: IEEE. **Image Analysis and Signal Processing, 2009. IASP 2009. International Conference on**. [S.l.], 2009. p. 172–175.
- WEI, K.; HUANG, J.; FU, S. A survey of e-commerce recommender systems. In: IEEE. **Service Systems and Service Management, 2007 International Conference on**. [S.l.], 2007. p. 1–5.
- WIDISINGHE, A. et al. picseek: Collaborative filtering for context-based image recommendation. In: IEEE. **Information and Automation for Sustainability (ICIAFs), 2010 5th International Conference on**. [S.l.], 2010. p. 225–232.
- WOERNDL, W.; SCHUELLER, C.; WOJTECH, R. A hybrid recommender system for context-aware recommendations of mobile applications. In: IEEE. **Data Engineering Workshop, 2007 IEEE 23rd International Conference on**. [S.l.], 2007. p. 871–878.

XU, S.; JIANG, H.; LAU, F. Observing facial expressions and gaze positions for personalized webpage recommendation. In: ACM. **Proceedings of the 12th International Conference on Electronic Commerce: Roadmap for the Future of Electronic Business**. [S.l.], 2010. p. 78–87.

YAN, B. **Web Recommendation System with Image Retrieval**. Dissertação (Mestrado) — Uppsala University, Department of Information Technology, 2011.

YAO, L. et al. Recommending web services via combining collaborative filtering with content-based features. In: IEEE. **Web Services (ICWS), 2013 IEEE 20th International Conference on**. [S.l.], 2013. p. 42–49.

ZHANG, D.; LU, G. Review of shape representation and description techniques. **Pattern recognition**, Elsevier, v. 37, n. 1, p. 1–19, 2004.

ZHOU, D. et al. A study of recommending locations on location-based social network by collaborative filtering. In: KOSSEIM, L.; INKPEN, D. (Ed.). **Advances in Artificial Intelligence**. [S.l.]: Springer Berlin Heidelberg, 2012, (Lecture Notes in Computer Science, v. 7310). p. 255–266.