

# Capítulo 1

## Introdução

Ao longo dos últimos quarenta anos, com a difusão de tecnologias eletrônicas modernas, a palavra “digital” tem se tornado cada vez mais lugar comum. Ao longo desse período, a tecnologia baseada em máquinas de cálculo automatizado que ocupavam prédios inteiros e eram operadas por poucos (para fins restritos) foi permeando nossa cultura global e ganhando destaque em vários campos do conhecimento e da vida humanas. Com o passar do tempo, as suas aplicações foram ganhando horizontes, os computadores passaram a ser pessoais e cada vez mais amplamente utilizados, contudo, ainda tinham aplicações específicas em ambientes de pesquisa e para atividades profissionais, e eram restritos a mesas e ambientes reservados. Em paralelo, a indústria de jogos eletrônicos começa a florescer, se fazendo presente nas residências através de seus consoles. Todos aproveitavam a evolução tecnológica apresentada pelo advento dos transistores e microprocessadores.

Adiante mais alguns anos, e a internet aparece, os microprocessadores estão mais “micro” e mais “processadores”, os consoles de video-game mais elaborados e seus clientes cada vez mais exigentes. Os microprocessadores vão transitando de máquinas estáticas para aplicações móveis: celulares e câmeras digitais (para foto e vídeo) tornam-se quase onipresentes. Mídias digitais tornam-se tão importantes quanto protocolos de comunicação, redes sociais movimentam a opinião pública e servem de plataforma a revoluções [4]. Muitos cidadãos globais têm contas em várias redes sociais e máquinas portáteis em seus bolsos, prontas a fazer um vídeo ou uma foto e postá-los para apreciação popular. Ao que se faz possível, com alguns poucos cliques e nenhum custo com passagem ou hospedagem, termos acesso à realidade atualizada de diversas partes do globo, em todos os idiomas possíveis, a qualquer hora.

O volume de dados gerado é tão grande que plataformas como o Flickr (da americana Yahoo) abrigava, em 2011, mais de seis bilhões de imagens digitais [11], e o crescimento estimado para o crescimento de seu banco de imagens nos anos seguintes foi de um bilhão por ano — e essa é apenas uma de várias plataformas de hospedagem de publicação de mídias [3] [8]. Outro exemplo, para publicação de vídeos, é o Youtube, onde seus mais de um milhão de usuários únicos assistem mais de seis bilhões de horas de vídeo por mês. Outras plataformas de publicação de vídeo também têm números expressivos, como o caso do Vimeo, com mais de 200 *petabytes* de vídeos executados em 2012 [14]. Além de serviços de hospedagem de conteúdo disponíveis publicamente, mais recentemente surgiram serviços de *streaming* de filmes e séries, como o Netflix, já com mais de 40 milhões de usuários em 41 países, disponibilizando mais de um bilhão de horas de vídeo [10].

Esses números representam um desafio para a indústria, que se responsabiliza por receber, armazenar e distribuir esses dados sob demanda, para todo o globo.

Voltemos um pouco à história de alguns parágrafos acima. Outra entidade que acompanha esse desenvolvimento, e que o antecede em algumas décadas é a televisão, cujo pai aclamado

é Philo Farnsworth, entre tantos outros engenheiros e contribuições [13]. Desde 1929 existe programação regular para TV sendo projetada no espaço, proveniente de ambos os lados do Atlântico (os Estados Unidos e o Reino Unido começaram a produzir programação regularmente na segunda metade de 1929) [13].

Com o nascimento dessa tecnologia, começou-se a discutir a necessidade de avaliar a qualidade da imagem recebida em aparelhos de TV, e artigos como “*Television images: an analysis of their essential qualities*”, publicado por Jesty e Wintch, em 1937[9], aparecem. Winch, começa seu artigo de 1953 [15] com a afirmação (em tradução livre) “a adição de cores à televisão traz muitos novos problemas a um assunto já complexo”. E já em 1940, Peter Goldmark e John Dyer [6] apresentam quais características são mais importantes ao determinar-se a qualidade de uma imagem (para TV): definição, faixa de contraste, ângulo de visualização (**gradation?**), brilho, efeitos da frequência de varredura (*flickering*), distorção geométrica, tamanho, cor e ruído. Algumas dessas características viriam a se tornar objetos de estudo da AQI nos anos vindouros, bem como algumas delas seriam bases para o cálculo de métricas de qualidade de imagem (fotografia e vídeo) — falaremos de algumas delas nesse trabalho.

Dado que a parte visual de um vídeo é constituída de imagens paradas em sequência, a avaliação da qualidade de vídeo e de imagem andam entrelaçadas desde o início. Tanto o é, que a International Telecommunications Union — ITU (União Internacional de Telecomunicações, tradução livre) não faz distinção em seus documentos de padronização de qualidade entre vídeo e imagem [1]; e seu grupo especializado para esse fim é chamado *Video Quality Experts Group* (VQEG, Grupo de Peritos em Qualidade de Vídeo, tradução livre). Em [1], encontramos recomendações para avaliação de qualidade de vídeo que podem ser também aplicadas a imagens. Nesse trabalho seguiremos essas recomendações.

Vê-se que, com tanta demanda por imagem e vídeo, e tráfego destes, é necessário que se encontre uma forma de armazená-los eficientemente, acessá-los confiavelmente e garantir que o usuário final terá a qualidade esperada, ainda que compressões e descarte de informações sejam necessários. Enquanto alguém preocupado com a compressão de uma imagem se perguntaria “qual a menor quantidade de informação necessária para que se mantenha a completude da mensagem (no nosso caso, imagem)?”, um pesquisador atual de AQI se pergunta: “como aferir a qualidade de uma imagem digital? Dado que a necessidade de reduzir a quantidade de informação é premente, o que garante qualidade percebida?”

Essa é uma pergunta razoavelmente complexa, já que envolve conceitos abstratos e subjetivos. Uma boa imagem para uma aplicação não o é, necessariamente para outra. Um exemplo simples é a aquisição de vídeos de segurança em comparação com a aquisição de vídeo para entretenimento. No primeiro caso, a qualidade mínima e suficiente é aquela que garante a identificação de um possível infrator; na segunda as exigências são mais altas (ninguém sentiria prazer em assistir toda a saga de Star Wars em preto e branco e com todo o ruído e baixa resolução que são aceitáveis para câmeras de segurança). É claro, não se pode deixar de lado as considerações sobre custo-benefício: no primeiro caso a resolução tem que ser mínima e suficiente para a identificação de um eventual infrator, mas também tem que ocupar pouco espaço em disco, já que câmeras de segurança, em sua maioria, funcionam continuamente. Quanto a Star Wars, o tamanho é fixo, uma vez terminada a edição; e a garantia da qualidade do produto final significa aumento de lucros em bilheterias pelo mundo afora.

Tradicionalmente, existem duas formas de se construir algoritmos de AQI que possam determinar automaticamente a adequação de uma imagem para consumo humano [5]: algoritmos baseados no sistema visual humano (SVH) e algoritmos baseados em sessões de avaliação.

Algoritmos baseados em SVH levam em consideração a física de todo um sistema biológico, bem como a parte psicológica da percepção visual. Esse trabalho não tratará desse tipo de algoritmo e o leitor é direcionado aos trabalhos de [2] e [16] para maiores informações.

A segunda abordagem, que será alvo de estudo nesse trabalho, não aborda a psicofísica e se concentra em características da imagem que sejam relevantes. Para obter esses dados, sessões de avaliação são organizadas, onde pessoas são questionadas a respeito da qualidade de um conjunto de imagens. Após coletados os dados, o pesquisador tenta produzir um modelo que tenha como saída uma nota aproximada daquela dada pelos entrevistados. Obviamente, quanto menor o erro entre o sistema modelado e a avaliação subjetiva dos indivíduos entrevistados, melhor o modelo.

Aqui trabalharemos com duas bases de imagens distintas, proveniente de grupos de pesquisa independentes, que coletaram avaliação para as imagens constantes em suas bases de forma muito similar. Trataremos dessa similaridade, e eventual diferença no cap /ref. Ambos os grupos de pesquisa, seguem o que é de praxe na área, e tratam seus procedimentos de avaliação estatisticamente com muito critério, segundo as recomendações do ITU e o usualmente praticado na área.

O questionamento que propomos é justamente sobre as ferramentas estatísticas utilizadas no tratamento desses dados e na validação das métricas em uso. Entendemos que, dada a natureza dos experimentos de avaliação e a forma como as notas são atribuídas às imagens, ferramentas de estatística em classes seriam mais adequadas para a análise dos resultados, e consequente ponderação de qualidade das métricas desenvolvidas sobre esses dados. Atualmente, as ferramentas estatísticas utilizadas nesse contexto são as mesmas utilizadas para tratar dados intrinsecamente não-categóricos.

Esse documento está estruturado em X partes, resumidamente descritas abaixo:

**Introdução:** breve histórico da área de Avaliação de Qualidade de Imagem, breve descrição do objetivo do trabalho e descrição sucinta das ferramentas utilizadas

**Antecedentes:** detalhamos melhor nossa proposta de trabalho, situamos o leitor quanto às abordagens estatísticas utilizadas na área e nossas ferramentas de comparação.

**Procedimentos Experimentais:** apresentamos as bases de imagem que serão utilizadas, os métodos utilizados e resultados preliminares de comparação com a literatura existente.

**Análise Categórica:** aprofundamos a explicação das ferramentas utilizadas, aplicamos essas ferramentas aos dados apresentados nos capítulos anteriores e efetuamos comparações de sua validade.

**Conclusão:** considerações finais sobre o trabalho e possíveis caminhos a serem tomados a partir das conclusões apresentadas.

**Referências:** lista de documentos que serviram de base para a produção desta obra.



## Parte I

# Preparação da pesquisa



## Capítulo 2

# Preâmbulo Teórico

Para que possamos fazer uma análise comparativa entre duas estratégias estatísticas, temos que, antes, entender suas diferenças e similaridades, bem como as características dos dados com os quais estamos lidando.

Assim, este capítulo será dividido em duas grandes seções, uma destinada à apresentação das bases de imagens utilizadas, outra destinada à discussão sobre as qualidades dos dados nessas bases.

### 2.1 As bases de dados

Para o cumprimento desse trabalho, escolhemos duas bases de imagens diferentes, que chamaremos simplesmente Toyama e LIVE.

A base Toyama é, na verdade, chamada **IRCCyN/IVC-Toyama database (LCD)** e tem acesso franqueado no site [12] do IRCCyN (*Institut de Recherche en Communications et Cybernétique de Nantes*), da Universidade de Nantes, na França.

A base LIVE é, na verdade, chamada **LIVE Image Quality Assessment Database** e tem acesso também franqueado no site [7] do LIVE (*Laboratory for Image & Video Engineering*).

Ambas as bases possuem imagens originais (não degradadas) e um determinado número de imagens degradadas com diferentes tipos e graus de degradação; nosso trabalho se concentra na degradação do tipo JPEG, em todos os graus de degradação apresentados. A Tabela 2.1 apresenta as principais características de ambas as bases.

Tabela 2.1: Características das Imagens nas Bases de Dados

	<b>Toyama</b>	<b>LIVE</b>
Número Total de Imagens $T_i$	98	204
Número de Imagens de Referência $I_r$	14	29
Número de Imagens degradadas $I_d$	84	175
Resolução das Imagens na base	768x512 pixels	768x512 pixels
Profundidade de cor	24bits/pixel	24bit/pixel
Formato das imagens cedidas	BMP	BMP
Tipo de degradação	JPEG	JPEG
Graus de degradação aplicados	15, 20, 27, 37, 55, 79	não informado
Diversidade de graus de degradação	6 taxas	não informado
Faixa de valores de avaliação	[1, 5]	[0, 100]
Categorias de Qualidade	5	5
Sessões de Avaliação distintas	1	2

Tabela comparativa de ambas as bases utilizadas.  $T_i = I_r + I_d$





## Parte II

# Procedimento Experimental



# Parte III

## Resultados



# Referências Bibliográficas

- [1] Tutorial - objective perceptual assessment of video quality: Full reference television.
- [2] *Developing a new psychophysical experimental method to estimate image quality*, volume 4421, 2002.
- [3] 500px Inc. 500px. website, December 2013.
- [4] Adriana Carranca. Revolução digital: da página no facebook à queda de mubarak. news website, February 2011.
- [5] Damon M. Chandler. Seven challenges in image quality assessment: Past, present, and future research. *ISRN Signal Processing*, 2013:53, 2013.
- [6] Peter C. Goldmark and J.N. Dyer. Quality in television pictures. *Proceedings of the IRE*, 28(8):343–350, 1940.
- [7] L. Cormack H.R. Sheikh, Z.Wang and A.C. Bovik. Live image quality assessment database release 2",. official website, February 2012.
- [8] Instagram Inc. Instagram press, 16 billion images. official website, December 2013.
- [9] L.C. Jesty and G.T. Winch. Television images: An analysis of their essential qualities. *Tran. Illum. Eng.*, 2:316–334, 1937.
- [10] Netflix Inc. Investor relations. official website, December 2013.
- [11] N. Olivarez-Giles. Flickr reaches 6 billion photos uploaded. news website, August 2011.
- [12] S. Tourancheau, F. Autrusseau, Z.M.P. Sazzad, and Y. Horita. Impact of subjective dataset on the performance of image quality metrics. In *Image Processing, 2008. ICIP 2008. 15th IEEE International Conference on*, pages 365–368, 2008.
- [13] various. History of television. wikipedia, December 2013.
- [14] Vimeo Inc. Official year-end self-appraisal and profound introspection: 2012. official staff blog, December 2012.
- [15] G.T. Winch. Colour television: some subjective and objective aspects of colour rendering. *Electrical Engineers, Journal of the Institution of*, 1953(4):191–192, 1953.
- [16] Stefan Winkler. *Digital Video Quality – Vision Models and Metrics*. John Wiley & Sons, January 2005.