Neurocomputing 275 (2018) 366–376

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Just Noticeable Difference for natural images using RMS contrast and feed-back mechanism

Vinit Jakhetiya[a,*], Weisi Lin[c], Sunil Jaiswal[b], Ke Gu[d], Sharath Chandra Guntuku[c]

[a] Department of Computer Science Engineering, Bennett University, Greater Noida
[b] Hong Kong University of Science and Technology, Hong Kong
[c] School of Computer Engineering, Nanyang Technological University, 639798 Singapore
[d] Beijing Key Laboratory of Computational Intelligence and Intelligent System, Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China

ARTICLE INFO

Article history:
Received 20 November 2016
Revised 8 April 2017
Accepted 15 August 2017
Available online 7 September 2017

Keywords:
Just Noticeable Difference
Contrast sensitivity
RMS contrast
Feedback mechanism
Bilateral filter

ABSTRACT

Contrast Sensitivity (CS), Luminance Adaptation (LA) and Contrast Masking (CM) are important contributing factors for Just Noticeable Difference (JND) in images. Most of the existing pixel domain JND algorithms are based only on LA and CM. Research shows that the human vision depends significantly on CS, and an underlying assumption in the existing algorithms is that CS cannot be estimated in the pixel domain JND algorithms. However, in the case of natural images, this assumption is not true. Studies on human vision suggest that CS can be estimated using the Root Mean Square (RMS) contrast in the pixel domain. With this in perspective, we propose the first pixel-based JND algorithm that includes a very important component of the human vision, namely CS by measuring RMS contrast. This RMS contrast is combined with LA and CM to form a comprehensive pixel-domain model to efficiently estimate JND in the low frequency regions. For high frequency regions (edge and texture), a feedback mechanism is proposed to efficiently alleviate the over- and under-estimation of CM. To facilitate this, a prediction based algorithm is used to classify an image into low (smooth) and high frequency regions. This feed-back mechanism is based on the relationship between the CS and RMS contrast. Experiments validate that the proposed JND algorithm efficiently matches with human perception and produces significantly better results when compared to existing pixel domain JND algorithms.

1. Introduction

Just Noticeable Difference (JND) is a visibility threshold below which change in an image cannot be sensed by the human visual system. In the human vision (HV), perceived information highly depends upon signal characteristics such as spatial frequency and contrast of signal [1–5]. In general, a change in a signal which is imperceptible to 75% of viewers can be defined as the JND value for the corresponding signal.

JND profiles are used in several multimedia applications. Information which cannot be sensed by the eyes can be removed with the guidance of JND and does not require to be coded in a bit-stream or this removed information help to enhance the accuracy of the image quality assessment matrices. Therefore, JND profiles are extensively used for multimedia coding [6,8,14,18,20]. The authors of [16,17,25,26] estimated video and image quality, respectively, using JND profiles. The JND profiles are also used for watermarking [15,19]. In the same line, researchers have also used JND profiles to guide visual signal enhancement [13,21]. Recently, JND profiles received a lot of attention and these profiles are used in different multimedia applications, such as, Fang et al. [27] used important component of JND (namely contrast sensitivity) for the saliency detection and authors of [28,29] used it for the depth sensation enhancement. Interestingly, JND profiles also have the ability to guide the seam carving [30] and enhancement of backlight-scaled images [31]. Such widespread use of JND reveals the significance of developing more accurate models to enhance the accuracy of multimedia applications.

Contrast sensitivity (CS), Luminance Adaptation (LA) and Contrast Masking (CM) are important contributing factors for JND in images. LA is the ability of human vision to adapt with the change in luminance and estimation of the LA is based upon the psychological experiments [20], and CM, which refers to the visibility

* Corresponding author.
E-mail addresses: vinit.jakhetiya@bennett.edu.in (V. Jakhetiya), wslin@ntu.edu.sg (W. Lin), spjaiswal@connect.ust.hk (S. Jaiswal), guke@bjut.edu.cn (K. Gu), sharathg@sas.upenn.edu (S.C. Guntuku).

http://dx.doi.org/10.1016/j.neucom.2017.08.031
0925-2312/

Table 1
Comparison between existing pixel domain JND models [6–8,20].

Attribute	Chou et al. [20]	NAMM [8]	Liu et al. [7]	Wu et al. [6]
Luminance Adaptation or Texture Masking	Yes	No	No	No
Luminance Adaptation and Contrast Masking	No	Yes	Yes	Yes
Image classification	No	No	Yes	No
Contrast sensitivity	No	No	No	No
CM based upon edge strength	No	Yes	Yes	No
CM based upon prediction errors	No	No	No	Yes
Under and/or over-estimation of CM	Yes	Yes	Yes	Yes

reduction of one signal in the presence of other signals, is estimated based upon the edge and texture strength [6–8,20]. CS, on the other hand, gives an idea of our visual system's ability to differentiate an object, from its background. In a broad sense, JND algorithms can be divided into two categories: pixel domain and sub-band. As the HV highly depends upon CS, existing sub-band JND algorithms estimate CS using the Contrast Sensitivity Function (CSF) in sub-bands [22]. However, existing pixel domain JND estimation algorithms do not have such capability due to the lack of an effective way to account for CS in pixel domain. For this reason, many sub-band JND algorithms have been proposed in the literature, while only a few pixel domain JND algorithms have been proposed. In many applications (such as image and video coding [6–8], enhancement [13] and quality assessment [17]), a direct pixel domain JND estimator is preferred. The importance of CS can be seen by the fact that model in [24] has to convert a pixel domain model into the sub-band, in order to consider CS. Therefore in this paper, we concentrate on pixel domain JND estimation which includes CS.

1.1. Related work

As aforementioned, a few pixel domain JND models have been previously proposed in the literature [6–8,20]. In general, HV is sensitive to the slowly varying regions, such as smooth regions and regions with weak edges, and small changes in these regions can be easily identified by HV. While HV is less sensitive to the high frequency regions and changes made in these regions are being unnoticed by HV. With this view, in existing JND algorithms, the main components for JND estimation are LA and CM [8] for low and high spatial frequency regions, respectively. The first attempt for pixel domain JND estimation was made in [20], in which JND threshold was estimated using either LA or contrast (texture) masking. In the same line, Yang et al. [8] proposed a JND model called the Non-linear Additivity Model for Masking (NAMM), which considers both components. Liu et al. [7] proposed an algorithm, in which an image is decomposed so that edge masking and texture masking are calculated separately for the edge and texture regions. Recently, Wu et al. [6] proposed a JND algorithm based upon the free-energy principle. In this model, prediction errors are used as the CM for the JND estimation for textural regions, while for the other regions, JND is estimated using the existing NAMM [8]. In general, these algorithms are not specifically designed for any applications. However, authors of these algorithms have applied corresponding models to remove the perceptual redundancy of images and/or videos for the coding purpose [6–8]. The main attributes of these existing pixel domain JND models [6–8,20] are summarized in Table 1.

From Table 1, one can observe that existing JND algorithms have two major issues:

(1) All the above described pixel domain JND algorithms [6–8,20] only use LA and CM for JND estimation. As such, these algorithms cannot include the effect of CS, which in turn leads to visible artefacts in images, especially in the smooth regions.

(2) Furthermore, in high spatial frequency areas (edge and texture), CM estimation in the existing algorithms is only based upon edge and texture strength [7,8,20] or prediction errors [6], which may lead to under- or over-estimation [22] of the CM. The changes made in the signal guided by the over estimated CM can be easily sensed by human visual system [1,2,22]. In these situations, inaccurate JND estimation can severely affect the efficiency of the multimedia applications.

To overcome the above-mentioned problems associated with the existing pixel domain JND algorithms, we propose a comprehensive and efficient pixel domain JND algorithm in which, we merge the effect of CS (by measuring RMS contrast) with LA and CM for estimating JND. We also propose a novel feedback mechanism, which efficiently alleviates the over- and under- estimation of the CM in high spatial frequency regions (such as edge and texture). The decomposition of an image into smooth, edge and texture regions is based on prediction errors of the input image. From experiments, it is validated that the proposed algorithm produces significantly better results as compared to existing pixel domain JND algorithms and efficiently matches with the human perception.

In short, the main contribution of this paper is to propose the first pixel domain JND algorithm to include the most important component of the HV, namely CS, which is combined with LA and CM for accurate JND estimation. This accurate estimation of JND can help to enhance the efficiency of the several multimedia applications.

The rest of the paper is organized as follows. Section 2 describes both the proposed decomposition of an image into smooth, edge and texture regions using prediction errors, and the JND algorithm. The comparison of the proposed algorithm with the existing state-of-the-art algorithms and results of subjective tests are provided in Section 3, and the discussions and concluding remarks are given in Sections 4 and 5, respectively.

2. Proposed contrast sensitivity and feedback mechanism based JND algorithm

In the proposed algorithm, we try to build a computational JND model which efficiently matches with human perception [1–5]. The relationship between the CS and spatial frequency is represented using the parabolic curve (CSF) [1,2]. In general for natural images, the HV is more sensitive towards the low spatial frequency regions when compared to high spatial frequency regions [1–8]. Therefore, JND should be higher in the high frequency regions as compared to low frequency regions [6–8]. Accordingly, we decompose an image into low and high frequency areas, based upon the prediction errors of the input image. For the low frequency areas, we propose a new algorithm which takes into account the effect of LA, CM and CS (RMS contrast), while for high frequency regions we propose to have a feed-back mechanism based upon the RMS contrast, which can efficiently control the CM.