



تخمین کمترین تفاوت قابل درک با استفاده از برجستگی بصری در تصاویر

فائزه نعمتی خلیلآباد^۱, هادی هادیزاده^۲, عباس ابراهیمی مقدم^۳ و مرتضی خادمی درح^{۴*}

^{۱,۲,۳}دانشکده مهندسی، دانشگاه فردوسی مشهد، مشهد، ایران

^۴دانشکده مهندسی، دانشگاه صنعتی قوچان، قوچان، ایران

چکیده

به علت وجود برخی محدودیت‌های فیزیولوژیکی و فیزیکی مختلف در مغز و چشم، دستگاه بینایی انسان (HVS) قادر به درک برخی تغییرات سیگنال بصری که دامنه آن‌ها از یک حد آستانه مشخص (موسوم به آستانه JND) پایین‌تر باشند، نیست. در بیش تر پژوهش‌های موجود جهت تخمین آستانه JND، حساسیت HVS در تمام صحنه یکسان در نظر گرفته شده و تأثیرات توجه بصری (VA) ناشی از برجستگی بصری (VS) در این پژوهش‌ها لحاظ نشده است. مطالعات مختلف نشان داده‌اند که حساسیت بصری در نواحی برجسته که توجه بصری بیشتری را جلب می‌کنند بیشتر بوده و لذا در آن نقاط آستانه JND پایین‌تر است و بالعکس. در این مقاله مدلی محاسباتی برای تخمین JND پیشنهاد می‌شود که از رابطه بین JND و برجستگی بصری برای بهبود تخمین آستانه JND استفاده می‌کند. این مدل با استفاده از یک مدل JND یکنواخت کارآمد و با به کارگیری یکتابع مدولاسیون غیر خطی مناسب، آستانه‌های JND پیکسل‌های مختلف در یک تصویر را با توجه به برجستگی بصری آن‌ها بهبود می‌دهد. تعیین پارامترهای تابع غیرخطی مدولاسیون در قالب یک مسئله بهینه‌سازی، مدل‌سازی می‌شود که حل آن منجر به یافتن مدل JND بهبود یافته می‌شود. کلید کارآمدی روش پیشنهادی به کارگیری سازوکاری است که منجر به استفاده کارآمدتر از برجستگی بصری می‌شود. آزمایش‌های انجام‌گرفته نشان دهنده برتری قابل ملاحظه روش پیشنهادی نسبت به روش‌های مشابه موجود است.

وازگان کلیدی: برجستگی بصری (VS)، توجه بصری (VA)، دستگاه بینایی مغز (HVS)، کمترین تفاوت قابل درک (JND).

Just Noticeable Difference Estimation Using Visual Saliency in Images

Faezeh Nemati Khalil Abad¹, Hadi Hadizadeh², Abbas Ebrahimi Moghadam³ & Morteza Khademi Darah^{4*}

^{1,3,4}Image Processing Laboratory, Faculty of Engineering, Ferdowsi University of Mashhad, Mashhad, Iran

²Image Processing Laboratory, Faculty of Electrical Engineering, Quchan University of Technology, Quchan, Iran

Abstract

Due to some physiological and physical limitations in the brain and the eye, the human visual system (HVS) is unable to perceive some changes in the visual signal whose range is lower than a certain threshold so-called just-noticeable distortion (JND) threshold. Visual attention (VA) provides a mechanism for selection of particular aspects of a visual scene so as to reduce the computational load on the brain. According to the current knowledge, it is believed that VA is driven by "visual saliency". In a visual scene, a region is said to be visually salient if it possess certain characteristics, which make it stand out from its surrounding regions

* Corresponding author

*نویسنده عهده‌دار مکاتبات

سال ۱۳۹۹ شماره ۲ پیاپی ۴۴

• تاریخ ارسال مقاله: ۱۳۹۷/۰۶/۲۲ • تاریخ پذیرش: ۱۳۹۸/۰۶/۱۱ • تاریخ انتشار: ۱۳۹۹/۰۶/۲۴ • نوع مطالعه: پژوهشی



and draw our attention to it. In most existing researches for estimating the JND threshold, the sensitivity of the HVS has been considered the same throughout the scene and the effects of visual attention (caused by visual saliency) which have been ignored. Several studies have shown that in salient areas that attract more visual attention, visual sensitivity is higher, and therefore the JND thresholds are lower in those points and vice versa. In other words, visual saliency modulates JND thresholds. Therefore, considering the effects of visual saliency on the JND threshold seems not only logical but also necessary. In this paper, we present an improved non-uniform model for estimating the JND threshold of images by considering the mechanism of visual attention and taking advantage of visual saliency that leads to non-uniformity of importance of different parts of an image. The proposed model, which has the ability to use any existing uniform JND model, improves the JND threshold of different pixels in an image according to the visual saliency and by using a non-linear modulation function. Obtaining the parameters of the nonlinear function through an optimization procedure leads to an improved JND model. What makes the proposed model efficient, both in terms of computational simplicity, accuracy, and applicability, are: choosing nonlinear modulation function with minimum computational complexity, choosing appropriate JND base model based on simplicity and accuracy and also Computational model for estimating visual saliency that accurately determines salient areas. Finally, determine the Efficient cost function and solve it by determining the appropriate objective Image Quality Assessment. To evaluate the proposed model, a set of objective and subjective experiments were performed on 10 selected images from the MIT database. For subjective experiment, A Two Alternative Forced Choice (2AFC) method was used to compare subjective image quality and for objective experiment SSIM and IWSIM was used. The obtained experimental results demonstrated that in subjective experiment the proposed model achieves significant superiority than other existing models and in objective experiment, on average, outperforms the compared models. The computational complexity of proposed model is also analyzed and shows that it has faster speed than compared models.

Keywords: Visual Saliency (VS), Visual Attention (VA), Human Visual System (HVS), Just Noticeable Difference (JND).

- CSF³] [8] به شدت نور (اندازه‌گیری می‌شود) یک تابع میان‌گذر است؛ یعنی مقدار حساسیت در فرکانس‌های میانی بالاتر از حساسیت در فرکانس‌های پایین و به خصوص فرکانس‌های بالا است.
 - (ب) نقاب‌گذاری روشنایی⁴ [9] که بیان می‌کند میزان درک تغییرات کوچک در نواحی خیلی روشن و خیلی تیره بیشتر از میزان درک تغییرات در نواحی با سطوح روشنایی میانه است. بهطور کلی، پدیده نقاب‌گذاری به معنای کاهش قابلیت دیدن یک سیگنال در حضور سیگنال دیگر در نزدیکی مکانی، زمانی و یا فرکانسی (طیفی)، آن است [1].
 - (ج) نقاب‌گذاری شدت نوری (بافتی)⁵ [8-11] که بیان گر کاهش حساسیت بصری در یک ناحیه از تصویر به دلیل افزایش غیرمکنواخت سطح روشنایی در آن ناحیه است.
 - (د) نقاب‌گذاری الگویی⁶ که توسط ابهام ساختاری⁷ و پیچیدگی الگو⁸ سنجیده می‌شود [12-13]. هر چه میزان پیچیدگی در یک الگو بیشتر باشد، نقاب‌های بصری قوی‌تری خواهد داشت.
- از دیگر محدودیتهای HVS توان محاسباتی آن

۱- مقدمه

کمترین تفاوت قابل درک (JND¹) در یک تصویر، آستانه‌ای است که چنانچه دامنه تغییرات یک سیگنال بصری از آن حد کمتر باشد آن تغییرات توسط دستگاه بینایی انسان (HVS²) قابل درک است. این آستانه بدليل برخی محدودیتهای فیزیکی HVS به وجود می‌آید ([1,2]) و از آنجا که JND توصیف‌کننده رؤیت‌پذیری تغییرات در دستگاه بینایی انسان است، مدل‌های محاسباتی مختلف آن برای تخمین افزونگی در فشرده‌سازی و ارسال تصاویر/ ویدئو [3-5]، پنهان‌سازی اطلاعات [6]، ارزیابی کیفیت تصویر [7] و غیره به کار می‌رود.

در طول دهه اخیر روش‌های مختلف متعددی برای تعیین آستانه JND هم در حوزه پیکسلی و هم در حوزه زیرباند ارائه شده است که در آن‌ها سعی شده تا عوامل مختلف تأثیرگذار بر آستانه JND مشخص شوند. از جمله عوامل مهم تأثیرگذار بر آستانه JND می‌توان به چهار عامل اصلی زیر اشاره کرد.

(الف) میزان حساسیت به شدت نور که توسط تابع حساسیت

این مقاله از پایان‌نامه کارشناسی ارشد در دانشگاه فردوسی با همین عنوان استخراج شده است.

¹ Just Noticeable Difference (JND)

² Human Visual System (HVS)

فصلنامه

پژوهش‌های
دانشگاه فردوسی
مشهد



آمده و آستانه‌های JND تخمین‌زده شده مبتنی بر نقشه برجستگی، توسط مجموعه‌ای از توابع مدولاسیون خطی^۵، تغییر مقیاس می‌شوند. در این روش پارامترهای توابع مدولاسیون به طور آزمایشی و تجربی طوری تنظیم می‌شوند تا با نتایج تجربی موجود در [۱۷] هماهنگ باشد. مدل مشابه جدیدتر و بهتری که از اثرات کنترلی و هدایتی توجه بصري می‌شود بهبود آستانه‌های JND استفاده کرده مدل پیشنهادی در [۲۰-۲۱] است. در این مدل که در حوزه DCT بلوکی انجام گرفته، آستانه JND هر بلوک بهوسیله دو تابع غیرخطی اشباع‌شونده که ورودی آن‌ها برجستگی‌های بصري هر بلوک است، بهبود داده می‌شود. پارامترهای این تابع غیرخطی از طریق حل یک مسأله بهینه‌سازی به دست می‌آید. همان‌طور که دیده می‌شود، این مدل همانند مدل [۱۹] در حوزه DCT بلوکی انجام گرفته، ولی از دو لحظه با مدل [۱۹] متفاوت است. نخست این که توابع مدولاسیون استفاده شده در مدل [۱۹] توابعی خطی هستند؛ در حالی که این توابع در مدل پیشنهادی در [۲۰-۲۱] توابعی غیرخطی اشباع‌شونده هستند. دوم این که پارامترهای توابع مدولاسیون استفاده شده در مدل [۱۹] به طور آزمایشی و تجربی تنظیم می‌شوند؛ در حالی که در مدل پیشنهادی در [۲۰-۲۱] پارامترهای یادشده به طور نظاممند و تطبیقی از طریق حل یک مسأله بهینه‌سازی به دست می‌آیند. در همین اواخر مدلی دیگر برای تعیین JND در حوزه پیکسل در [۲۲] ارائه شده که در آن از اطلاعات برجستگی بصري یک تصویر جهت تخمین بهتر آستانه‌های JND به‌گونه‌ای استفاده کرده است که فاصله ادراکی^۶ بین تصویر اولیه و تصویر JND که به آن نویه تزریق شده است، کمینه شود. بدین منظور ابتدا یک فاصله ادراکی وزن‌دارشده توسط برجستگی و مشتق‌پذیر در حوزه لاپلاسین نرمالیزه شده^۷ تعریف می‌شود؛ سپس با یک فرآیند بهینه‌سازی با هدف کمینه کردن فاصله ادراکی بین تصویر JND نویه‌ای شده با تصویر اولیه به تخمین بهتر آستانه JND می‌پردازد. در [۵۴] مدل JND در حوزه تبدیل موجک^۸ ارائه شده که این مدل مشابه با مدل ارائه شده در [۲۰] است.

در این مقاله، با درنظر گرفتن مطالب یادشده بالا، مدلی جدید برای بهبود آستانه‌های JND در حوزه پیکسل پیشنهاد کردۀایم. مدل‌های حوزه پیکسلی برای محاسبه JND نیاز به انتقال به حوزه زیرباند را ندارند؛ لذا برای

است که باعث می‌شود پردازش همزمان داده‌های بصري ورودی به HVS که بسیار حجم هستند بسیار دشوار شود. ازین‌رو، به‌منظور کاهش حجم اطلاعات مورد نیاز برای پردازش در مغز، HVS از سازوکاری موسوم به سازوکار توجه بصري (VA^۱) استفاده می‌کند. مکانیزم توجه بصري باعث می‌شود که در هنگام تماشای یک تصویر، تنها بر بخش کوچکی از تصویر که از اهمیت حیاتی و کاربردی بیشتری برخوردار است، تمرکز شود و جزئیات در بخش‌های دیگر تصویر مورد توجه قرار نگیرد. در طول سالیان گذشته، به‌دلیل کاربرد و اهمیت بالای توجه بصري، یافتن مدل محاسباتی برای آن یکی از موضوعات مهم پژوهشی بوده که توجه فراوانی را به خود جلب کرده و چندین مدل محاسباتی نیز برای آن ارائه شده است [۱۴-۱۶]. مدل‌های محاسباتی ارائه شده برای تصاویر ثابت، با تولید یک نقشه موسوم به نقشه برجستگی^۲، نقاط برجسته تصویر را که به‌دلیل دارابودن مشخصه‌های خاصی باعث تمایز از مناطق اطراف و جلب توجه می‌شود، تعیین می‌کند. به عبارت بهتر، برجستگی بصري (VS^۳) با سازوکار غیریکنواخت‌سازی اهمیت نقاط مختلف یک تصویر به تعیین نقاط حساس و مهم آن پرداخته و از این طریق آستانه JND را تحت تأثیر قرار داده و آن را کنترل و اصلاح می‌کند؛ ازین‌رو، به‌منظور تخمین دقیق‌تر آستانه JND، لحاظ‌کردن اثرات هدایتی برجستگی بصري بر روی این آستانه، نه تنها منطقی، بلکه لازم و ضروری به نظر می‌رسد [۱۷-۱۹].

در بیشتر مدل‌های تخمین JND موجود، حساسیت HVS در تمام صحنه یکسان فرض شده و تأثیرات توجه بصري در این مدل‌ها لحاظ نشده است [۹],[۲۳],[۵۳]. حال آنکه مشخص شده که حساسیت بصري در نواحی برجسته‌تر بیشتر بوده و لذا در آن نقاط آستانه JND پایین‌تر است و بالعکس، هر چه برجستگی بصري ناحیه‌ای کمتر باشد، حساسیت بصري آن کمتر و آستانه JND آن بالاتر است. برخلاف وجود مدل‌های مختلف متعدد برای تخمین آستانه JND، با این حال مدل‌های JND بسیار اندکی وجود دارد که تأثیر برجستگی بصري را بر روی آستانه‌های رویت‌پذیری در نظر گرفته‌اند [۱۸-۲۲]. [۵۴]. یکی از مدل‌های اولیه قابل توجه در زمینه JND هدایت شده بهوسیله برجستگی بصري، مدل پیشنهادی در [۱۹] است که بر پایه یک مدل JND ترکیبی در حوزه تبدیل کسینوسی گسسته (DCT^۴) به دست

¹ Visual Attention (VA)

² Saliency Map

³ Visual Saliency (VS)

⁴ Discrete Cosine Transform (DCT)

⁵ Linear Modulation Functions

⁶ Perceptual Distance

⁷ Normalized Laplacian

⁸ Wavelet transform

برجستگی‌های بصری لحاظ نشده است که این باعث فاصله داشتن آستانه JND به دست آمده از مقدار واقعی و بهینه آن می‌شود. از طرف دیگر، همان‌طور که پیش‌تر اشاره شد، بهدلیل سازوکار توجه بصری، مقدار آستانه‌های بصری در نواحی برجسته از آستانه‌های بصری در نواحی غیربرجسته پایین‌تر هستند. به همین منظور، در این مقاله با لحاظ‌کردن توجه بصری و مقدار برjستگی تصویر در نقاط مختلف، یک مدل JND غیریکنواخت (از لحاظ توجه بصری) پیشنهاد می‌شود که در آن آستانه‌های JND توسط مفهوم برjستگی بصری تنظیم و اصلاح می‌شود.

شکل (۱) نمایی کلی از مدل پیشنهادی است که در آن I_{ref} تصویر اولیه است که می‌خواهیم برای آن نقشه JND را تعیین کنیم، JND نقشه JND ناشی از هر مدل یکنواخت دلخواه موجود است که قرار است با بهره‌بردن از برjستگی‌های بصری آن را بهبود دهیم^۱. نقشه JND بهبودیافته ناشی از مدل غیریکنواخت پیشنهادی است. مدل غیریکنواخت پیشنهادی قابلیت استفاده از هر مدل یکنواخت دلخواه موجود را به عنوان مدل پایه اولیه (JND) دارد؛ ولی از آنجا که مدل‌های حوزه پیکسلی برای محاسبه JND نیاز به انتقال به حوزه زیرباند را ندارند و از لحاظ محاسباتی ساده‌تر هستند، روش پیشنهادی نیز از یک مدل یکنواخت که در همین اواخر در حوزه پیکسلی ارائه شده و نتایج خوبی به همراه داشته [۲۳] به عنوان مدل پایه اولیه (یعنی JND) استفاده کرده و آن را بهبود می‌دهد.

همان‌طور که در شکل (۱) مشاهده می‌شود JND با استفاده از یک تابع (ذاتاً غیرخطی) از JND و نقشه برjستگی‌های تصویر (SM) ورودی به دست می‌آید. در حقیقت IJND گونه هدایت (کنترل) شده JND با تابع مدولاسیون برjستگی است که این هدایت‌شدگی با توجه به برjستگی‌های بصری تصویر به صورت زیر محاسبه می‌شود:

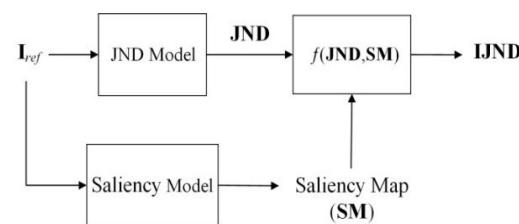
$$IJND = f(JND, SM), \quad (1)$$

که در آن، SM نقشه برjستگی تصویر ورودی است. بدیهی است با توجه به رابطه (۱)، هر چه مدل پایه اولیه (یعنی JND) و نقشه برjستگی (یعنی SM) و همچنین تابع IJND مدولاسیون برjستگی مناسب‌تر استفاده شود بهتری حاصل می‌شود. در این مقاله، برای مدل پایه اولیه JND، همان‌طور که اشاره شد، از مدل مناسب ارائه شده در [۲۳] استفاده می‌شود و برای نقشه برjستگی SM نیز با

^۱ Improved JND (IJND)

تعیین آستانه JND در برنامه‌های کاربردی از لحاظ محاسباتی ساده‌تر هستند. نظر به این‌که امروزه پردازش داده‌های تصویری حجمی یکی از چالش‌های مهم پردازش تصویر است، ارائه مدل‌ها و روش‌های پردازش تصویر که از کمینه پیچیدگی محاسباتی برخوردار باشند از اهمیت فوق العاده‌ای برخوردارند. همچنین وجه تمایز مدل پیشنهادی نسبت به بیش‌تر مدل‌های JND است. برjستگی‌های بصری برای بهبود تخمین آستانه JND است. مدل پیشنهادی که قابلیت اعمال روی هر مدل JND دلخواه موجود را دارد با بهره‌بردن از برjستگی‌های بصری و به کارگیری یکتابع مدولاسیون غیرخطی که پارامترهای آن در قالب یک مسأله بهینه‌سازی مدل‌سازی می‌شود به بهبود بیشتر تخمین آستانه‌های JND نواحی مختلف نسبت به مدل‌های مشابه موجود می‌پردازد. آنچه که باعث برتری روش پیشنهادی نسبت به روش‌های مشابه موجود می‌شود استفاده کارآمدتر از برjستگی‌های بصری جهت بهبود آستانه JND است که این کار از طریق استفاده از یک متربیک بصری کارآمدتر و در نتیجه تعریف تابع هزینه بهتر و حل مناسب‌تر مسأله بهینه‌سازی تشکیل‌شده انجام می‌گیرد. همچنین جهت ارزیابی روش پیشنهادی، آزمایش‌های ذهنی ارائه می‌شود که نشان‌دهنده برتری و کارآمدی قابل ملاحظه روش پیشنهادی است.

ساختار مقاله به این ترتیب است که در بخش دوم روش پیشنهادی بطور کامل بیان می‌شود. ارزیابی‌ها و نتایج تجربی در بخش سوم ارائه می‌شود. درنهایت، در بخش چهارم، به جمع‌بندی و نتیجه‌گیری مطالب پرداخته می‌شود.



(شکل-۱): نمودار بلوکی کلی مدل پیشنهادی

(Figure-1): Block diagram of the proposed model.

۲- روش پیشنهادی

همان‌طور که بیان شد، بیش‌تر قریب به اتفاق مدل‌های JND موجود از لحاظ توجه بصری، یکنواخت هستند، یعنی با فرض یکسان‌بودن میزان اهمیت بصری در تمامی صحنه به دست آمده‌اند و در آن‌ها تأثیر توجه بصری و

D (I_{ref}, I_{noisy}) یک معیار ارزیابی عینی کیفیت بصری تصویر با مرجع⁴ (FR-IQA) است [25-28] که از نظر بصری نشان‌دهنده فاصله بین تصویر مرجع اصلی (عینی I_{ref}) و نسخه نوفه‌ای شده تصویر اصلی (عینی I_{noisy}) است. هر چه مقدار این عبارت کمتر باشد، کیفیت بصری تصویر نوفه‌ای تولیدشده به کیفیت بصری تصویر خام اصلی نزدیک‌تر خواهد بود. با توجه به کارآمدی معیار شباهت ساختاری وزن دارشده می‌تنی بر محتوای اطلاعاتی⁵ (IW-SSIM) [28] و همچنین امکان لحاظ‌کردن برجستگی بصری در آن، در این مقاله ما از IW-SSIM بعنوان معیار ارزیابی عینی کیفیت بصری تصویر استفاده می‌کنیم؛ درنهایت در رابطه (۳)، λ یک پارامتر کنترلی جهت کنترل میزان اهمیت و مشارکت دو جمله MSE و D در تابع هزینه تعریف شده، است. شکل (۲) نمای همراه با جزئیات کامل‌تر نمودار بلوکی پیشنهادی را نشان می‌دهد که جزئیات عملکرد آن در ادامه تشریح می‌شود.

۱-۲- تولید نقشه JND یکنواخت اولیه

همان‌طور که بیان شد، در این مقاله ما قصد داریم که مدل JND غیریکنواختی را ارائه دهیم که با درنظرگرفتن یک مدل JND یکنواخت دلخواه موجود به عنوان مدل پایه اولیه (JND) و بهره‌بردن از برجستگی‌های بصری، آن را بهبود دهد. با لحاظ کردن دو عامل سادگی محاسباتی و کارآبی، در این مقاله ما از مدل JND یکنواختی که در همین اواخر در [23] ارائه شده به عنوان مدل پایه اولیه (عینی JND) استفاده کرده و آن را بهبود می‌دهیم. مدل یادشده به دلیل پیاده‌سازی در حوزه پیکسل، نسبت به مدل‌های حوزه زیرباند از لحاظ محاسباتی ساده‌تر و همچنین با توجه به نتایج ارائه شده در [23] از کارآبی بهتری هم نسبت به روش‌های مشابه موجود برخوردار است. با توجه به این که هر چه میزان پیچیدگی یک الگو بیشتر باشد، نقاب‌های بصری قوی تری وجود خواهد داشت، در [23] ابتدا تنوع زاویه‌ای و جهت‌های مختلف پیکسل‌ها را به عنوان معیاری جهت تعیین پیچیدگی یک الگو تعريف می‌کند؛ سپس با ترکیب پیچیدگی الگو و نقاب‌گذاری وضوح، یک تابع نقاب‌گذاری مکانی⁶ جدید به صورت زیر تعريف می‌کند:

$$M_S(x) = \max \{M_P(x), M_C(x)\}, \quad (4)$$

⁴ Full-Reference Image Quality Assessment

⁵ Information Content Weighted Structural Similarity (IW- SSIM)

⁶ Spatial Masking

توجه به نتایج بدست‌آمده، مدل مناسب و کارآمد ارائه شده در [24] به کار گرفته می‌شود؛ لذا با توجه به رابطه (۱)، تنها عامل باقیمانده جهت رسیدن به IJND مناسب، تخمین مناسب تابع مدولاسیون f است که این کار در این مقاله با در نظر گرفتن یک فرم کلی برای تابع غیرخطی f و سپس تعیین پارامترهای اصلی آن از طریق یک فرآیند بهینه‌سازی انجام می‌گیرد. برای این منظور و پرهیز از اضافه‌شدن پیچیدگی و بار محاسباتی، در اینجا از تابع خطاطا^۱ (به اختصار crf) که غیرخطی بوده و فقط یک ضریب مجھول دارد برای تابع مدولاسیون برجستگی مطابق زیر استفاده می‌شود:

$$h(S(x), \alpha) = 1 + \operatorname{erf}(\alpha S(x)) \\ = 1 + \frac{2}{\sqrt{\pi}} \int_0^{\alpha S(x)} e^{-t^2} dt, \quad (2)$$

که در آن، S(x) برجستگی نرمالیزه شده در پیکسل x و crf(α(S(x))) مقدار تابع خطاطا در پیکسل x است که با یک ضریب α تعریف می‌شود. با تغییر پارامتر تابع خطاطا، شکل آن می‌تواند از خطی تا غیرخطی تغییر کند. برای تعیین پارامتر تابع مدولاسیون (عینی ضریب α) که در واقع تعیین کننده میزان غیرخطی بودن آن است، ابتدا تابع هزینه^۲ زیر را تعریف می‌کنیم:

$$J(I_{noisy}(\alpha)|I_{ref}S) = D(I_{ref}, I_{noisy}(\alpha)) - \lambda \operatorname{MSE}(I_{ref}, I_{noisy}(\alpha)), \quad (3)$$

که در آن، J تابع هزینه، MSE³ میانگین مربعات خطاطا (یا انرژی نوفه)، I_{ref} تصویر مرجع اصلی است که می‌خواهیم برای آن نقشه JND را تعیین کنیم، I_{noisy} نسخه نوفه‌ای شده تصویر اصلی است که این افزودن نوفه با توجه به نقشه JND بهبودیافته (عینی IJND) انجام می‌گیرد، به عبارت دیگر، قدر مطلق دامنه نوفه اضافه شده در هر نقطه از تصویر برابر با اندازه آستانه IJND در آن نقطه است. انرژی نوفه افزوده شده به تصویر اصلی یعنی انرژی نوفه موجود در I_{noisy} با جمله دوم در رابطه (۳) به دست می‌آید. در مسئله مورد نظر ما می‌خواهیم کیفیت بصری I_{ref} تا حد امکان مشابه باشد؛ چون نوفه اضافه شده توسط مدل JND هدایت و کنترل می‌شود و لذا اگر مدل پیشنهادی به خوبی عمل کند نوفه اضافه شده نباید قابل درک و مشاهده باشد. برای اندازه‌گیری میزان مرئی بودن نوفه اضافه شده، از جمله نخست رابطه (۳) استفاده می‌کنیم. در این جمله،

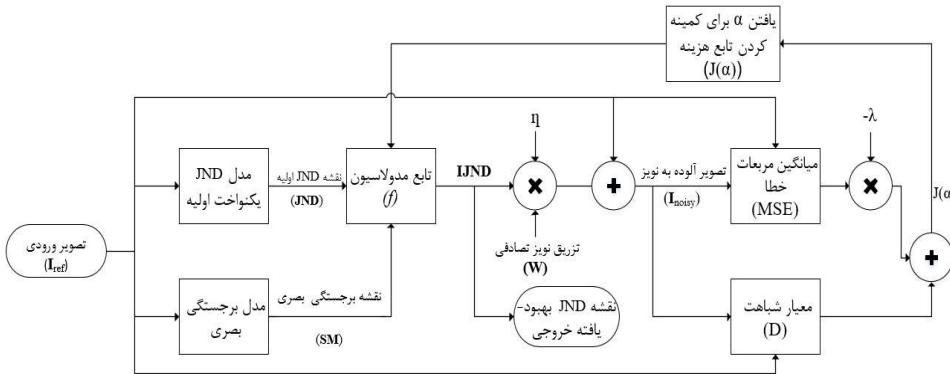
¹ Error Function

² Cost Function

³ Mean Square Error (MSE)

$$\text{تعیین می‌شود:} \\ \text{T}_{\text{JND}}(x) = L_{\text{A}}(x) + M_{\text{S}}(x) - 0.3 \min\{L_{\text{A}}(x), M_{\text{S}}(x)\}. \quad (5)$$

که در آن، M_{S} اثر نقاب‌گذاری مکانی کل، M_{C} اثر نقاب‌گذاری وضوح و M_{P} اثر نقاب‌گذاری الگویی است. در نهایت با ترکیب نقاب‌گذاری مکانی فوق و تطبیق روشناهی آستانه JND پیکسل x ($\text{T}_{\text{JND}}(x)$) توسط رابطه زیر



(شکل-۲): نمودار بلوکی مفصل‌تر مدل پیشنهادی برای بهبود تخمین آستانه JND با استفاده از بر جستگی بصری.

(Figure-2): More detailed block diagram of the proposed model to improve the JND threshold estimation using visual saliency

نقشه خیرگی^۲ چشم توسط یک شبکه پیش‌خور^۳ محاسبه و یک مدل جدید ارائه شده که توانایی پیش‌بینی دقیق نقشه بر جستگی را از طریق ترکیب و یکی کردن سازوکارهای توجه عصبی دارد. هسته اصلی راه حل پیشنهادی، یک حافظه طولانی کوتاه‌مدت کانولوشنی^۴ است که روی بر جسته‌ترین نواحی تصویر ورودی متمرکز شده و مکرر، نقشه بر جستگی پیش‌بینی شده را اصلاح می‌کند.

۲-۳- تولید نقشه JND غیریکنواخت بهبود یافته

پس از آشناییدن با نحوه تولید نقشه‌های JND یکنواخت اولیه و بر جستگی بصری، حال به تشریح چگونگی استفاده از نقشه بر جستگی جهت بهبود مدل JND اولیه و تولید نقشه JND غیریکنواخت بهبود یافته می‌پردازیم. با توجه به روابط (۱) و (۲)، مقدار JND غیریکنواخت پیکسل x (یعنی $\text{IJND}(x)$) به صورت زیر محاسبه می‌شود:

$$\text{IJND}(x) = \text{JND}(x) \odot h(S(x), \alpha), \quad (6)$$

که در آن، (۱) نشان‌دهنده ضرب پیکسلی و (۲) مقدار JND یکنواخت پیکسل x است؛ درنتیجه، در این قسمت، هدف، یافتن نقشه JND دقیق‌تری از تصویر ورودی است؛ بهنحوی که با تزریق نویه بر اساس مقادیر تعیین شده در این

² Gaze Map

³ Feed-Forward Network

⁴ Convolutional Long Short-Term Memory (LSTM)

۲-۲- تولید نقشه بر جستگی بصری

برای تخمین بر جستگی بصری و تولید نقشه بر جستگی جهت استفاده در روش پیشنهادی، دو راهیافت متفاوت زیر قابل استفاده است:

(الف) استفاده از یک مدل محاسباتی مشهور برای تخمین بر جستگی بصری، همانند مدل‌های ارائه شده در [14]، [24]، [29] و [30]. گرچه مدل‌های محاسباتی موجود به طور معمول چندان دقیق نیستند، اما به علت جنبه‌های کاربردی و تجاری‌سازی، استفاده از این مدل‌های محاسباتی امری لازم و اجتناب‌ناپذیر است.

(ب) استفاده از یک پایگاه داده تصاویر (همانند [31]) است که نقشه‌های بر جستگی بصری آن‌ها توسط آزمایش‌های ردیابی چشم^۱ به دست آمده است. مزیت استفاده از این روش این است که با استفاده از این روش، می‌توان به مقدار واقعی بر جستگی‌های بصری تصاویر مختلف دست یافت.

با توجه به مزایا و معایب دو راهیافت بالا، ما سعی کردیم از مدلی جهت تولید نقشه بر جستگی استفاده کنیم که تا حد امکان مزایای هر دو راهیافت را داشته باشد؛ لذا با توجه به نتایج ارائه شده در [24] و همچنین [32]، ما از مدل محاسباتی [24] که هم دارای دقت بسیار خوبی است و هم قابلیت اعمال بر روی هر تصویر دلخواه را دارد (جنبه کاربردی و تجاری‌سازی) استفاده کردیم. درواقع، در [24]

¹ Eye-Tracking



كه در آن، IWSSIM مقدار شباهت ساختاري وزن دار بين تصاوير I_{noisy} و I_{ref} است. صرفنظر از کارآمدی اثبات شده معيار IW-SSIM [28] آنچه که باعث مناسب بودن استفاده اين معيار در روش ما مى شود، وجود وزن هاي مهم و مؤثر در تعين شباهت ساختاري بين تصاوير I_{noisy} و I_{ref} است که در اين مقاله سازوکار تعين وزن با استفاده از نقشه برجستگي است.

برای حل مسأله بهينه سازی (۸)، از روش گوسی-نيوتون تصادفي استفاده شده که به منظور يافتن دقیق تر نقطه بهینه، فرآيند بهينه سازی يكصدبار تکرار و در هر بار اجرا، نقطه شروع فرآيند بهينه سازی به صورت تصادفي تعين شده است. همچنان جهت رسیدن به بهترین جواب در روش پيشنهادی، مقادير $\lambda=50$ و $\eta=1$ استفاده شده اند که از طریق تجربی به دست آمده است.

۳- نتایج تجربی

در اين مقاله به منظور ارزیابی روش پیشنهادی، آزمایش های مختلفی انجام شده است. به منظور پیاده سازی روش پیشنهادی، از بانک های اطلاعاتی رديابی چشم [31] استفاده شده است. اين بانک اطلاعاتی حاوی 100^3 تصویر RGB با رزولوشن 1024×768 پیکسل است. اين بانک اطلاعاتی حاوی تصاویر متنوعی از تصاویر چهره تا تصاویر مناظر و چشم انداز های طبیعی است که ده تصویر از اين مجموعه جهت انجام آزمایش ها انتخاب شده اند (تصویر نشان داده شده در شکل (۳)).

از آن جايی که بهترین روش ارزیابی کيفيت تصاویر، روش مبتنی بر آزمایش های ذهنی^۲ است، به منظور ارزیابی دقت و عملکرد واقعی مدل پیشنهادی در تخمین JND برای يك تصویر، از آزمایش های ذهنی استفاده می کنیم. برای اين منظور استفاده از روش مشهور^۳ 2AFC [33] پیشنهاد می شود. در اين آزمایش، مدل پیشنهادی ما با مدل JND يکنواخت مقایسه می شود. در 2AFC، از شرکت كنندگان خواسته می شود که بين دو تصویر (يکی تصویر نفاهی شده با استفاده از مدل JND يکنواخت و دیگری تصویر نفاهی شده با استفاده از مدل پیشنهادی)، تصویری که کيفيت بهتری را دارد، انتخاب کنند. برای شرکت در اين آزمون، از نوزده شرکت كننده خواسته شد که به دو تصویر کنار هم (با فاصله

نقشه، بتوان تصویر نفاهی قابل قبولی به لحاظ کيفيت بصري نسبت به تصویر اولیه به دست آورد. برای تولید تصویر نفاهی شده I_{noisy} با استفاده از تصویر مرجع I_{ref} و بر اساس نقشه JND بهبود یافته (يعني IJND) از رابطه زير استفاده می کنیم:

$$I_{noisy} = I_{ref} + \eta IJND \odot W, \quad (7)$$

كه در آن، W ماتريسي با ابعاد يكسان با تصویر مرجع I_{ref} و با درایه های تصادفي $+1$ و -1 است و η پارامتری کنترلي جهت تنظيم ميزان انرژي نفه تزريقي که توسط PSNR^۱ و يا MSE محاسبه می شود، است. دليل استفاده از ماتريس تصادفي W ، جلوگيري از ايجاد اثرات و الگوهای ساختگی و تصنعي در تصویر نويزي حاصله است. توجه داريد که در هنگام مقایسه عملکرد و کيفيت مدل های مختلف JND با يكديگر، ابتدا يك سطح نفه مشخص را به وسیله يکي از مدل ها به يك تصویر نمونه تزريقي می کنیم؛ سپس همان نفه (يعني ماتريس W) را توسط مدل دیگر به همان تصویر اضافه کرده و با استفاده از پارامتر η ، انرژي نفه در تصویر دوم را به گونه های تغيير می دهيم تا PSNR و يا MSE هر دو تصویر نفه های حاصله با يكديگر برابر شود. تحت انرژي يكسان نفه، مدلی که دارای کيفيت ادرارکی بهتری باشد، مدل بهتری خواهد بود.

همان طور که اشاره شد، ما به دنبال آن هستيم تا با داشتن نقشه برجستگي، مدل JND بهبود یافته ای (IJND) را به دست آوريم بطور يکه کيفيت بصري مشابه I_{ref} I_{noisy} باشد. با توجه به رابطه (۳)، اين معادل آن است که $D(I_{ref}, I_{noisy})$ تا حد امكان کوچک باشد و در عين حال برای به دست آوردن بزرگ ترین سطح آستانه JND ممکن، انرژي نفه ($MSE(I_{ref}, I_{noisy})$)^(a) می بايست تا حد امكان بيشينه شود. اين خواسته با کمينه سازیتابع هزينه (۳) برآورده می شود؛ لذا، ما با يك مسأله بهينه سازی به صورت زير مواجه هستيم:

$$\begin{aligned} & \arg \min_{\alpha} J(I_{noisy}(\alpha) | I_{ref}, S) = \\ & \arg \min_{\alpha} \left\{ D(I_{ref}, I_{noisy}(\alpha)) \right. \\ & \quad \left. - \lambda MSE(I_{ref}, I_{noisy}(\alpha)) \right\} \end{aligned} \quad (8)$$

همان طور که پيشتر بيان شد، برای متريک D از معيار کارآمد IW-SSIM [28] مطابق زير استفاده می کنیم:

$$D(I_{ref}, I_{noisy}(\alpha)) = 1 - IWSSIM, \quad (9)$$

¹ Peak Signal To Noise Ratio (PSNR)

² Subjective Test

³ Two Alternative Forced Choice



(شکل-۳): مجموعه تصاویر انتخابی جهت انجام آزمایش‌ها، از بالا

راست تا پایین چپ به ترتیب از ۱ تا ۱۰ شماره‌گذاری شده‌اند.

(Figure-3): Image set for subjective viewing test, from top right to bottom left, they are named as 1-10.

در جدول (۱)، بر روی ده تصویر انتخابی شکل (۳)، آزمایش‌های ذهنی صورت گرفته شده است. تصاویر در سه گروه مورد سنجش قرار گرفته‌اند. تصاویر گروه A، تصاویری هستند که در آن تابع هزینه کمترین مقدار را دارد. از آنجا که کمترین مقدار تابع هزینه باعث متفاوت بودن مقدار PSNR تصاویر می‌شود در گروه B تصاویری با مقدار PSNR یکسان (در حدود ۲۸ دسی‌بل) ارزیابی شده و در گروه C، تصاویر در حدود ۲۱ دسی‌بل) مورد سنجش قرار گرفته است.

در جدول (۱)، در ستون دوم مجموع تعداد رأی‌هایی که تصویر نوفه‌ای شده توسط مدل JND یکنواخت را دارای کیفیت بهتر می‌دانند و در ستون سوم، مجموع رأی‌هایی که تصویر نوفه‌ای تولیدشده توسط مدل پیشنهادی دارای کیفیت بهتر می‌دانند، نوشته شده است. برای سنجش اهمیت آماری نتایج به دست آمده، از یک آزمون χ^2 ^۱ [25] استفاده شد که خروجی آن موسوم به مقدار p است [۲۵]. فرض پوج^۲ در این آزمون آن است که هیچ تفاوتی بین تصویر نوفه‌ای شده توسط مدل JND یکنواخت و تصویر نوفه‌ای تولیدشده به وسیله مدل پیشنهادی وجود ندارد. به عبارت دیگر فرض می‌شود که اگر فرض پوج صحیح باشد،

عمودی یکسان و یک فاصله افقی یک سانتی‌متری) بر روی یک پس‌زمینه نیمه خاکستری نگاه کنند. هر تصویر به مدت پنج ثانیه نمایش داده شد و پس از آن، پس‌زمینه نیمه‌خاکستری به مدت پنج ثانیه نمایش داده شد. از شرکت‌کنندگان خواسته شد در طول بازه زمانی پنج ثانیه دوم تصویر با کیفیت بهتر را، صرف نظر از میزان اطمینان از تصمیم خود، در پاسخ نامه مشخص کنند (تصویر سمت چپ و یا سمت راست). شرکت‌کنندگان در مورد نحوه قرارگیری تصاویر اطلاعی نداشتند و نمی‌دانستند که کدام تصویر، تصویر نوفه‌ای شده با استفاده از مدل پیشنهادی است و کدام یک، تصویر نوفه‌ای شده با استفاده از مدل JND یکنواخت است. همچنین به منظور مقابله با سوگیری افراد، به صورت اتفاقی، نیمی از تصاویر نوفه‌ای شده مدل پیشنهادی در سمت راست صفحه و نیمی دیگر در سمت چپ صفحه قرار داده شد و هر آزمایش با جایه‌جاکردن تصاویر سمت چپ و راست دو بار انجام گرفت. آزمایش در اتاق ساکت با نوزده شرکت‌کننده انجام گرفت. تمامی شرکت‌کنندگان بینایی سالم و یا بینایی تصحیح شده داشتند. در این آزمایش از صفحه نمایش ۱۵/۶ اینچ ASUS با وضوح ۳۸۴۰×۲۱۶۰ استفاده شد. سطح روشنایی مانیتور ۵۰٪ روشنایی اتاق در حدود ۳۲۰-۳۹۰ LUX و فاصله بین صفحه نمایش و شرکت‌کنندگان، پنجاه سانتیمتر تنظیم شد. قبل از شروع آزمایش به هر شرکت‌کننده نحوه انجام آزمایش توضیح داده شد. مدت زمان تقریبی هر آزمایش برای هر شرکت‌کننده در حدود دوازده دقیقه بود.

^۱ Two-Sided Pearson's Chi-Square (χ^2) Test^۲ Null Hypothesis

(جدول-۲): مقایسه مدل JND غیریکنواخت پیشنهادی با مدل IWSSIM JND یکنواخت اولیه بر اساس مقادیر SSIM و SSIM (Table-2): Comparing the proposed non-uniform JND model and the primary uniform JND model, based on the SSIM and IWSSIM values.

IWSSIM		SSIM		شماره تصویر
غیرJND مدل	یکنواخت اولیه	غیرJND مدل	یکنواخت پیشنهادی اولیه	
0.9819	0.97	0.85	0.85	A1
0.984	0.9751	0.84	0.84	A2
0.9836	0.9719	0.769	0.7759	A3
0.9854	0.9735	0.9139	0.9115	A4
0.9859	0.9774	0.8452	0.8449	A5
0.9856	0.9744	0.7943	0.8069	A6
0.986	0.977	0.6085	0.6159	A7
0.9765	0.9598	0.918	0.9152	A8
0.9876	0.9723	0.8267	0.8423	A9
0.9837	0.978	0.8878	0.8917	A10
0.9796	0.9642	0.8419	0.8316	B1
0.9691	0.9519	0.7515	0.7538	B2
0.9671	0.943	0.6849	0.6851	B3
0.9669	0.9399	0.8639	0.8639	B4
0.9718	0.9532	0.7716	0.7776	B5
0.9702	0.9454	0.694	0.698	B6
0.9666	0.9398	0.5693	0.5664	B7
0.9687	0.9452	0.8954	0.8889	B8
0.9876	0.9723	0.665	0.6874	B9
0.9777	0.9698	0.8605	0.8637	B10
0.9203	0.8666	0.5914	0.5718	C1
0.9014	0.8441	0.5172	0.5094	C2
0.903	0.8389	0.5307	0.5235	C3
0.9272	0.8732	0.77	0.7651	C4
0.9239	0.871	0.6083	0.6156	C5
0.9265	0.8658	0.5278	0.8273	C6
0.9227	0.8729	0.5325	0.5275	C7
0.9221	0.869	0.7844	0.7733	C8
0.9669	0.9286	0.4907	0.5117	C9
0.9304	0.908	0.702	0.7042	C10

جدول (۲) به بررسی اهمیت آزمایش‌های عینی^۱ در ارزیابی کیفی تصاویر می‌پردازد. برای بررسی و مقایسه کیفیت تصاویر نویزی شده توسط مدل JND یکنواخت و مدل JND پیشنهادی از دو معیار ارزیابی کیفیت بصری SSIM و IWSSIM استفاده شده است. در تمام نقاط تصویر دارای یک میزان اهمیت است؛ لذا با بررسی ستون دوم و سوم جدول (۲) مشاهده می‌شود که در بعضی از تصاویر، تصویر نویزهای شده بر اساس مدل JND یکنواخت دارای مقدار SSIM بیشتری نسبت به مدل پیشنهادی است. با استفاده از SSIM وزن دار (یعنی IWSSIM) نقاط مختلف تصویر به صورت غیر یکنواخت وزن دهنده می‌شود و ما با استفاده از این مدل، می‌توانیم وزن نقاط مختلف را با توجه به میزان برجستگی آنها لاحظ کنیم. با بررسی ستون چهارم و پنجم جدول (۲) مشاهده می‌شود که در تمامی تصاویر،

^۱ Objective Test

نیمی از شرکت‌کنندگان به تصویر اصلی و نیمی دیگر به تصویر حاصل از مدل پیشنهادی رأی داده‌اند. احتمال اینکه فرض پوج برقرار باشد، در اصطلاح مقدار p نامیده می‌شود. به عنوان یک قاعده کلی، وقتی $p < 0.05$ است، فرضیه پوج اولیه رد می‌شود و این بدین معنی است که یک تصویر تعداد قابل توجهی از آرا را در مقایسه با تصویر دیگر به دست آورده است و درنتیجه دارای کیفیت بهتری است. با توجه به جدول (۱)، در می‌یابیم که در تمامی تصاویر، مقدار p کمتر از ۰/۰۵ بوده است. به عبارتی دیگر، این بیان‌گر آن است که تصویر نویزهای شده بر اساس مدل پیشنهادی ما در مقایسه با تصویر نویزهای شده بر اساس مدل JND یکنواخت از نظر کیفی برتری مطلق دارد.

(جدول-۱): مقایسه مدل JND غیریکنواخت پیشنهادی با مدل JND یکنواخت اولیه بر اساس تعداد آرای جمع آوری شده از نوزده شرکت‌کننده (هر شرکت‌کننده دو بار).

(Table-1): Comparing the proposed non-uniform JND model and the primary uniform JND model, based on the number of votes collected from 19 participants (each participant twice).

p-value	غیریکنواخت JND مدل پیشنهادی	یکنواخت اولیه	شماره تصویر
≈0	38	0	A1
≈0	38	0	A2
≈0	38	0	A3
5.6E-06	33	5	A4
5.6E-06	33	5	A5
≈0	38	0	A6
≈0	38	0	A7
2.47E-05	32	6	A8
1.14E-06	34	4	A9
0	38	0	A10
≈0	38	0	B1
≈0	38	0	B2
≈0	38	0	B3
3.5E-08	36	2	B4
2.1E-07	35	3	B5
≈0	38	0	B6
≈0	38	0	B7
≈0	38	0	B8
≈0	38	0	B9
≈0	38	0	B10
≈0	38	0	C1
≈0	38	0	C2
≈0	38	0	C3
≈0	38	0	C4
1E-08	37	1	C5
7E-10	38	0	C6
≈0	38	0	C7
≈0	38	0	C8
≈0	38	0	C9
≈0	38	0	C10
3.4E-09	1114	26	مجموع

در جدول (۳) همانند جدول (۲) اهمیت آزمایش‌های عینی در ارزیابی کیفیت تصاویر مورد ارزیابی قرار گرفته و مقدار IWSSIM مدل JND غیریکنواخت پیشنهادی و مدل غیریکنواخت ارائه شده در [20] معروف به مدل SJND مورد مقایسه قرار گرفته است. بررسی و مقایسه مقادیر IWSSIM مدل پیشنهادی و مدل SJND، نشان‌دهنده برتری مدل پیشنهادی در بیشتر تصاویر آزمایش است.

پیچیدگی محاسباتی مدل JND غیریکنواخت پیشنهادی در جدول (۴) مورد ارزیابی قرار گرفته است. 7 بدين منظور سه مدل JND برروي لپ تاپ (core-Intel i7-6700HQ CPU 2.60GHZ- 8GIG RAM) مورد ارزیابی قرار گرفتند.

(جدول-۴): پیچیدگی محاسباتی سه مدل JND

(Table-4): Computational complexity analysis of different JND models.

SJND	غیریکنواخت پیشنهادی JND	یکنواخت JND (pattern complexity)	مدل
3.57s	1.32s	1.08s	زمان

در شکل (۳) تصویر یازدهم (تصویر گل با ابعاد $619 \times 585 \times 3$) برای انجام آزمایش مورد استفاده قرار گرفته است. همان‌طور که مشاهده می‌شود زمان اجرای مدل JND غیریکنواخت پیشنهادی خیلی بهتر از مدل SJND است؛ اما در مقایسه با مدل JND یکنواخت زمان اجرا اندکی افزایش یافته است. این افزایش زمان به این دلیل است که مدل JND پیشنهادی بهبود یافته و غیریکنواخت شده مدل JND یکنواخت است.

شکل (۴) به مقایسه شهودی و بصری مدل JND پیشنهادی و مدل JND یکنواخت می‌پردازد. نوفه تزریقی به دو مدل یکسان و الگوی آن (ماتریس W) نیز یکسان است تا مقایسه دو مدل در شرایط یکسان انجام شود. مقایسه دو ستون سمت چپ نشان می‌دهد که کیفیت مدل پیشنهادی بهتر است.

۴- نتیجه‌گیری

برخلاف پژوهش‌های فراوانی که در زمینه تخمین JND انجام شده است، تعداد بسیار اندکی از مقالات، اثر بر جستگی بصری را در تعیین JND و بهبود آن در نظر گرفته‌اند. در این مقاله، با درنظر گرفتن سازوکار توجه بصری و بهره‌گرفتن از بر جستگی بصری که منجر به غیریکنواخت شدن اهمیت نقاط مختلف یک تصویر می‌شود، مدل بهبود یافته غیریکنواختی برای تخمین آستانه JND تصاویر ارائه کردیم.

تصویر نوفه‌ای شده بر اساس مدل پیشنهادی، کیفیت بهتری نسبت به تصویر نوفه‌ای شده بر اساس مدل JND یکنواخت دارد. در آزمایش‌های گروه C، به خاطر بهتر مشاهده شدن تفاوت کیفی بصری دو مدل، میزان نوفه تزریقی به تصویر IWSSIM در این حالت، برتری مدل پیشنهادی را نسبت به مدل JND یکنواخت بیشتر نمایان می‌کند. گفتنی است که هر الگوریتم ارزیابی کیفیت تصاویر^۱ به طور معمول از دو مرحله معیار کیفیت محلی و ادغام^۲ تشکیل شده است. معیار ارزیابی IWSSIM بدليل استفاده از یک فرآیند ادغام کارا که توسط سازوکار وزن دهی بر اساس محتوای اطلاعاتی انجام می‌گیرد نسبت به معیار SSIM در ارزیابی تصاویر کارآمدتر است. در این مقاله، بهمنظور بهبود فرآیند ادغام، بر جستگی بصری نقاط در سازوکار وزن دهی لحاظ شده و باعث بهبود چشم‌گیر در نتایج حاصله شده است.

(جدول-۳): مقایسه مدل JND غیریکنواخت پیشنهادی با مدل

IWSSIM غیریکنواخت SJND (۲۰) بر اساس مقادیر

(Table-3): Comparing the proposed non-uniform JND model and the SJND model, based on the IWSSIM values.

IWSSIM	شماره تصویر	غیریکنواخت پیشنهادی JND مدل	SJND
0.9812	A1	0.9776	0.984
0.984	A2	0.9741	0.984
0.984	A3	0.973	0.985
0.985	A4	0.999	0.9858
0.9858	A5	0.9888	0.9855
0.9855	A6	0.9737	0.9858
0.9858	A7	0.9619	0.9769
0.9769	A8	0.9721	0.9878
0.9878	A9	0.9731	0.9839
0.9839	A10	0.9883	0.9752
0.9752	B1	0.9574	0.9611
0.9611	B2	0.9507	0.9646
0.9646	B3	0.9581	0.9644
0.9644	B4	0.9576	0.9655
0.9655	B5	0.9601	0.9642
0.9642	B6	0.9587	0.9585
0.9585	B7	0.9405	0.9644
0.9644	B8	0.9561	0.9641
0.9641	B9	0.9584	0.9739
0.9739	B10	0.9699	0.9146
0.9146	C1	0.8778	0.9611
0.9611	C2	0.9507	0.8799
0.8799	C3	0.8772	0.9218
0.9218	C4	0.9165	0.9059
0.9059	C5	0.8884	0.8941
0.8941	C6	0.8776	0.8868
0.8868	C7	0.835	0.8935
0.8935	C8	0.8661	0.896
0.896	C9	0.8762	0.9146
0.9146	C10	0.9237	

¹ Image Quality Assessment (IQA)² Pooling

اگر چه ایده کلی استفاده از برجستگی‌های بصری جهت بهبود آستانه JND در قبل به طور بسیار محدود و ساده مطرح و استفاده شده اما به دلیل عدم پرداخت شایسته به آن، ظرفیت بالقوه بالای آن در جهت بهبود تخمین JND مغفول مانده است. در این مقاله تلاش شد تا با نگاهی دقیق‌تر و عمیق‌تر این مسأله بررسی شود و با استفاده کاراتر از برجستگی‌های بصری جهت بهبود آستانه JND، مدلی کارآمدتر نسبت به مدل‌های اندک موجود ارائه شود.



(شکل-۴): سه مثال بصری جهت مقایسه کیفیت بصری مدل پیشنهادی با مدل JND یکنواخت در PSNR یکسان از راست به چپ: تصاویر اصلی، نقشه‌های برجستگی بصری، تصاویر تولیدشده بر اساس مدل JND یکنواخت و تصاویر تولیدشده بر اساس مدل پیشنهادی. در ردیف اول و ستون‌های سوم و چهارم مقدار PSNR برابر 21.84 dB است، در ردیف دوم و ستون‌های سوم و چهارم مقدار PSNR برابر 21.93 dB است، در ردیف سوم و ستون‌های سوم و چهارم مقدار PSNR برابر 21.61 dB است.

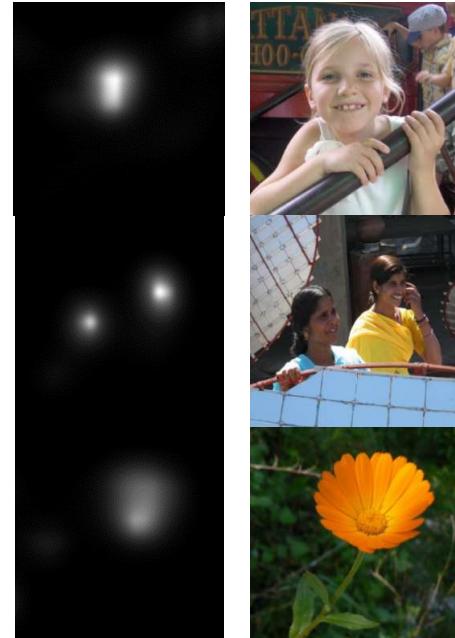
(Figure-4-4): Three visual examples for comparing the visual quality of the proposed model with the uniform JND model in the same PSNR. From right to left: Original images, visual saliency maps, generated images based on the uniform JND model and generated images based on the proposed model. In the first row and in the third and fourth columns, the PSNR value is equal to 21.84 dB, in the second row and the third and fourth columns, the PSNR value is equal to 21.93 dB, in the third row and the third and fourth columns, the PSNR value is equal to 21.61 dB.

همچنین از مزایای روش پیشنهادی، تعیین پارامتر تابع مدولاسیون غیرخطی به صورت تطبیقی بر اساس محتوای تصویر است.

۲- انتخاب مناسب هم مدل JND پایه که در عین سادگی (به دلیل پیاده‌سازی در حوزه پیکسلی)، از دقت خوبی نیز برخوردار است و هم مدل محاسباتی تخمین برجستگی بصری که در عین دقیق‌بودن، کاربردی نیز است و با شرایط واقعی مطابقت دارد.

۳- با توجه به مسأله بهینه‌سازی تشکیل شده و نقش

مدل پیشنهادی که قابلیت اعمال روی هر مدل JND یکنواخت دلخواه را دارد با به کارگیری یک تابع مدولاسیون غیرخطی، آستانه‌های JND نواحی مختلف تصویر را با توجه به برجستگی بصری آنها بهبود می‌دهد. تعیین پارامتر تابع غیرخطی مدولاسیون در قالب یک مسأله بهینه‌سازی مدل‌سازی و کمینه‌سازی تابع هزینه پیشنهادی منجر به یافتن مدل JND غیریکنواخت بهبود یافته شد. تابع هزینه پیشنهادی به گونه‌ای طراحی شده است که با توجه به میزان کیفیت بصری، نو费 بیشینه‌ای را به تصویر تزریق می‌کند.



مجموعه عواملی که باعث کارآمدی روش پیشنهادی چه از نظر سادگی محاسباتی، چه از لحاظ دقت و چه کاربردی بودن آن می‌شود عبارتند از:

- انتخاب مناسب تابع غیرخطی مدولاسیون که هم شرایط مطلوب مورد نیاز را برآورده می‌سازد و هم از سادگی لازم و کمینه پیچیدگی برخوردار است. این امر هم در ساده‌تر شدن مسأله بهینه‌سازی تشکیل شده (و درنتیجه حل آن) و هم در استفاده مناسب از برجستگی بصری و رسیدن به نتایج مطلوب تأثیرگذار است.

- Vis. Commun. Image Represent*, vol. 19, no. 1, pp. 30–41, 2008.
- [10] C. H. Chou and Y. C. Li, “A perceptually tuned subband image coder based on the measure of just-noticeable distortion profile,” *IEEE Trans. Circuits Syst. Video Technol*, vol. 5, no. 6, pp. 467–476, 1995.
- [11] Z. Wei and K. Ngan, “Spatio-temporal just noticeable distortion profile for grey scale image/video in dct domain,” *IEEE Trans. Circuits Syst. Video Technology*, vol. 19, no. 3, pp. 337–346, 2009.
- [12] J. Wu, W. Lin, G. Shi, X. Wang, and F. Li, “Pattern masking estimation in image with structural uncertainty,” *IEEE Trans. Image Process*, vol. 22, no. 12, pp. 4892–4904, 2013.
- [13] J. Wu, L. Li, W. Dong, G. Shi, W. Lin, C. J. Kuo, “Enhanced just noticeable difference model for images with pattern complexity,” *IEEE Trans. Image Process*, vol. 26, no. 6, pp. 2682–2693, 2017.
- [14] مهدی بنی‌طالبی دهکردی، عباس ابراهیمی مقدم، مرتضی خادمی، هادی هادی‌زاده. "تشخیص نقاط بررسی تصاویر با استفاده از نمونه‌برداری فشرده در حوزه موجک". فصلنامه پردازش علائم و داده‌ها. دوره ۱۶ شماره ۴، ۵۹-۷۲، ۱۳۹۸.
- [14] M. Banitalebi-Dehkordi, A. Ebrahimi-moghadam, M. Khademi, H. Hadizadch. “Compressed-Sampling-Based Image Saliency Detection in the Wavelet Domain”, *JSDP*, vol. 16 (4), pp. 59-72, 2020
- [15] L. Itti, G. Rees, and J. K. Tsotsos, *Neurobiology of Attention*. Academic Press, 2005.
- [16] A. Borji and L. Itti, “State-of-the-art in visual attention modeling,” *IEEE Trans. Pattern Anal. Mach. Intell*, vol. 35, no. 1, pp. 185–207, 2013.
- [17] L. Itti, J. Braun, C. Koch, “Modeling the modulatory effect of attention on human spatial vision,” *Advances in Neural Information Processing Systems (NIPS)*, MA, USA: MIT Press, vol. 14, pp. 1247–1254, 2002.
- [18] Z. Lu, W. Lin, X. Yang, E. Ong, and S. Yao, “Modeling visual attention’s modulatory aftereffects on visual sensitivity and quality evaluation,” *IEEE Trans. Image Process*, vol. 14, no. 11, pp. 1928–1942, 2005.
- [19] Y. Niu, M. Kyan, L. Ma, A. Beghdadi, S. Krishnan, “Visual saliency’s modulatory effect on just noticeable distortion profile and its application in image watermarking,” *Signal Process. Image Commun*, vol. 28, no. 8, pp. 917–928, 2013.

تعیین کننده متريک (I_{ref}, I_{noisy}) D در حل آن، انتخاب معيار کارآمد IWSSIM که هم بهطور مناسبی معکوس کننده ميزان شباهت‌های بصری است، و هم با لحاظ‌کردن اثر برجستگی بصری در مکانیزم وزن‌دهی، باعث استفاده کارآمدتر از برجستگی‌های بصری می‌شود و همچنین منجر به سخت‌شدن حل مسأله بهينه‌سازی تشکيل شده نمی‌شود، بسیار مهم و کلیدی است. درنهایت، کلیه آزمایش‌های ذهنی و عینی انجام شده جهت ارزیابی روش پیشنهادی که بر روی ده تصویر انتخابی از بانک اطلاعاتی MIT (شکل (۳)) انجام گرفته، نشان‌دهنده بهبودی قابل ملاحظه روش پیشنهادی در تخمین دقیق‌تر آستانه‌های JND است.

۵- مراجع

- [1] A. B. Watson, *Digital Images and Human Vision*. The MIT press, 1993.
- [2] F. A. A. Kingdom, *Psychophysics: A Practical Introduction*. Academic press, 2009.
- [3] X. K. Yang, W. S. Lin, Z. K. Lu, E. P. Ong, and S. S. Yao, “Just noticeable distortion model and its applications in video coding,” *Signal Process.: Image Community*, vol. 20, no. 7, pp. 662–680, 2005.
- [4] X. Yang, W. Lin, Z. Lu, E. Ong, S. Yao, “Motion-compensated residue pre-processing in video coding based on just-noticeable-distortion profile,” *IEEE Trans. Circuits Syst. Video Techno*, vol. 15, no. 6, pp. 742–752, 2005.
- [5] H. R. Wu, A. R. Reibman, W. Lin, F. Pereira, and S. S. Hemami, “Perceptual visual signal compression and transmission,” *Proceedings of The IEEE*, vol. 101, no. 9, pp. 2025–2043, 2013.
- [6] C. H. Chou and K. C. Liu, “A perceptually tuned watermarking scheme for color images,” *IEEE Trans. Image Process*, vol. 19, no. 11, pp. 2966 – 2982, 2010.
- [7] W. Lin and C. J. Kuo, “Perceptual visual quality metrics: A survey,” *J. Visual Communication and Image Representation*, vol. 22, no. 4, pp. 297–312, 2011.
- [8] A. Liu, W. Lin, M. Paul, C. Deng, and F. Zhang, “Just noticeable difference for images with decomposition model for separating edge and textured regions,” *IEEE Trans. Circuits Syst. Video Technolo*, vol. 20, no. 11, pp. 1648–1652, 2010.
- [9] X. Zhang, W. Lin, and P. Xue, “Just-noticeable difference estimation with pixels in images,” *J.*



- [34] M. Uzair, R. D. Dony, "Estimating just-noticeable distortion for images/videos in pixel domain", *IET Image Processing*, vol. 11, no. 8, pp. 559–567, 2017.
- [35] C. Wang, X. Han, W. Wan, J. Li, J. Sun, and M. Xu, "Visual saliency based just noticeable difference estimation in DWT domain," *Information*, vol. 9, no. 7, pp. 178, 2018.
- فائزه نعمتی** دانشجوی کارشناسی ارشد دانشگاه فردوسی است. وی مدرک کارشناسی خود را در رشته برق گرایش مخابرات در سال ۸۸ دریافت کرد. علایق پژوهشی ایشان پردازش سیگنال، بینایی و یادگیری ماشین است.
- نشانی رایانامه ایشان عبارت است از:
fa.nematykh@mail.um.ac.ir
- هادی هادیزاده** از بهمن ۱۳۹۲ عضو هیأت علمی گروه مهندسی برق-مخابرات دانشگاه صنعتی قوچان است. ایشان دکترا تحصیلی خود را در رشته مهندسی مخابرات از دانشگاه سایمون فریزر کانادا در سال ۱۳۹۲ اخذ کرد. زمینه‌های تخصصی ایشان عبارتند از پردازش و فشرده‌سازی تصاویر و ویدیوهای دیجیتال، شبکه‌های مخابراتی، بینایی و یادگیری ماشین، بازشناسی الگو و مدل‌سازی دستگاه بینایی مغز انسان.
- نشانی رایانامه ایشان عبارت است از:
h.hadizadeh@qiet.ac.ir
- عباس ابراهیمی** مقدم مدرک کارشناسی برق گرایش الکترونیک خود را از دانشگاه صنعتی شریف در سال ۱۳۷۰ و در سال ۱۳۷۴ مدرک کارشناسی ارشد را از دانشگاه صنعتی خواجه نصیر اخذ کرده است. ایشان مدرک دکترا خود را از دانشگاه McMaster کانادا دریافت کرده‌اند و از سال ۱۳۹۰ به عنوان استادیار در دانشگاه فردوسی مشهد فعالیت علمی می‌کنند. زمینه‌های پژوهشی مورد علاقه ایشان پردازش گفتار، پردازش تصویر و ویدیو، بینایی ماشین و پردازش سیگنال‌های حیاتی است.
- نشانی رایانامه ایشان عبارت است از:
a.ebrahimi@um.ac.ir
- [20] H. Hadizadeh, "A saliency-modulated just-noticeable-distortion model with non-linear saliency modulation functions," *Pattern Recognition Letters*, vol. 84, pp. 49–55, 2016.
- [21] H. Hadizadeh, "Energy-efficient images," *IEEE Trans. on Image Process*, vol. 26, no. 6, pp. 2882–2891, 2017.
- [22] H. Hadizadeh, A. Rajati, and I. V. Bajic, "Saliency-guided just noticeable distortion estimation using the normalized laplacian pyramid," *IEEE Signal Processing Letters*, vol. 24, 2017.
- [23] J. Wu, L. Li, W. Dong, G. Shi, W. Lin, C. J. Kuo, "Enhanced just noticeable difference model for images with pattern complexity," *IEEE Trans. Image Process*, vol. 26, no. 6, pp. 2682–2693, 2017.
- [24] M. Cornia, L. Baraldi, G. Serra, and R. Cucchiara, "Predicting human eye fixations via an LSTM-based saliency attentive model," [Online]. Available: <https://arxiv.org/abs/1611.09571>, 2017.
- [25] Z. Wang, A. C. Bovik, H. R. Sheikh, E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process*, vol. 13, no. 4, pp. 1–14, 2004.
- [26] L. Zhang, L. Zhang, X. Mou, D. Zhang, "FSTM: A feature similarity index for image quality assessment," *IEEE Trans. Image Process.*, vol. 20, no. 8, pp. 2378–2386, 2011.
- [27] L. Zhang, Y. Shen, H. Li, "VSI: A visual saliency-induced index for perceptual image quality assessment," *IEEE Trans. Image Process*, vol. 23, no. 10, pp. 4270–4281, 2014.
- [28] Z. Wang and Q. Li, "Information content weighting for perceptual image quality assessment," *IEEE Trans. Image Process*, vol. 20, no. 5, pp. 1185–1198, May 2011.
- [29] L. Zhang, Z. Gu, and H. Li, "SDSP: A novel saliency detection method by combining simple priors," *Proc. IEEE Int. Conf. Image Process*, pp. 171–175, Sep. 2013.
- [30] A. Borji, M.-Ming Cheng, H. Jiang, and J. Li, "Salient object detection: A benchmark," *IEEE Trans. on Image Process*, vol. 24, no. 12, pp. 5706–5722, 2015.
- [31] T. Judd, K. Ehinger, F. Durand, and A. Torralba, "Learning to predict where humans look," *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, pp. 2106–2113, 2009.
- [32] http://saliency.mit.edu/results_cat2000.html.
- [33] M. M. Taylor, C. D. Creelman, "PEST: efficient estimates on probability functions," *J. Acoust. Soc. Am.*, vol. 41, pp. 782–787, 1967.



فائزه نعمتی دانشجوی کارشناسی ارشد دانشگاه فردوسی است. وی مدرک کارشناسی خود را در رشته برق گرایش مخابرات در سال ۸۸ دریافت کرد. علایق پژوهشی ایشان پردازش سیگنال، بینایی و یادگیری ماشین است.

نشانی رایانامه ایشان عبارت است از:

fa.nematykh@mail.um.ac.ir



هادی هادیزاده از بهمن ۱۳۹۲ عضو هیأت علمی گروه مهندسی برق-مخابرات دانشگاه صنعتی قوچان است. ایشان دکترا تحصیلی خود را در رشته مهندسی مخابرات از دانشگاه سایمون فریزر کانادا در سال ۱۳۹۲ اخذ کرد. زمینه‌های تخصصی ایشان عبارتند از پردازش و فشرده‌سازی تصاویر و ویدیوهای دیجیتال، شبکه‌های مخابراتی، بینایی و یادگیری ماشین، بازشناسی الگو و مدل‌سازی دستگاه بینایی مغز انسان.

نشانی رایانامه ایشان عبارت است از:

h.hadizadeh@qiet.ac.ir



Abbas ابراهیمی مقدم مدرک کارشناسی برق گرایش الکترونیک خود را از دانشگاه صنعتی شریف در سال ۱۳۷۰ و در سال ۱۳۷۴ مدرک کارشناسی ارشد را از دانشگاه صنعتی خواجه نصیر اخذ کرده است. ایشان مدرک دکترا خود را از دانشگاه McMaster کانادا دریافت کرده‌اند و از سال ۱۳۹۰ به عنوان استادیار در دانشگاه فردوسی مشهد فعالیت علمی می‌کنند. زمینه‌های پژوهشی مورد علاقه ایشان پردازش گفتار، پردازش تصویر و ویدیو، بینایی ماشین و پردازش سیگنال‌های حیاتی است.

نشانی رایانامه ایشان عبارت است از:

a.ebrahimi@um.ac.ir

مرتضی خادمی تحصیلات خود را در مقاطع کارشناسی و کارشناسی ارشد مهندسی برق به ترتیب در سال‌های ۱۳۶۴ و ۱۳۶۶ در دانشگاه صنعتی اصفهان به پایان رسانده است. نامبرده



سال ۱۳۷۰ الی ۱۳۷۴ به عنوان عضو هیأت علمی (مربي) در دانشگاه فردوسی مشهد به کار مشغول بود. پس از آن به دوره دکترای مهندسی برق در دانشگاه ولونگونگ (استرالیا) وارد شده و در سال ۱۳۷۴ موفق به اخذ درجه دکترا در مهندسی برق از دانشگاه یادشده شد. دکتر خادمی از سال ۱۳۷۴ دوبلره در دانشکده مهندسی دانشگاه فردوسی مشهد مشغول به فعالیت شد و اینک نیز استاد این دانشکده است. زمینه‌های علمی مورد علاقه نامبرده شامل موضوعاتی مانند مخابرات ویدئویی، فشرده‌سازی ویدئو، پردازش تصویر، پردازش سیگنال‌های پزشکی و پنهان‌سازی اطلاعات در ویدئو است.

نشانی رایانمۀ ایشان عبارت است از:

khademi@um.ac.ir