



Semnan University

Journal of Modeling in Engineering

Journal homepage: <https://modelling.semnan.ac.ir/>

Research Article

Pyramid Image Fusion Based on Contourlet and Enhanced Structural Decomposition

Mojtaba Soleimani¹, Ali Aghagolzadeh^{2,*}, Mehdi Ezoji

1. PhD student of Faculty of Electrical and Computer Engineering Babol Noshirvani University of Technology, Babol, Iran
2. Professor of Faculty of Electrical and Computer Engineering Babol Noshirvani University of Technology, Babol, Iran
3. Associated professor of Faculty of Electrical and Computer Engineering Babol Noshirvani University of Technology, Babol, Iran

*Corresponding Author: Aghagol@nit.ac.ir

PAPER INFO

PAPER HISTORY:

Received: 04 September 2022

Revised: 31 December 2022

Accepted: 31 January 2023

Keywords:

Image Fusion,
Image Structural
Decomposition,
Contourlet Transform,
Pyramidal Structure,
Proposed Weighting
Factor.

ABSTRACT

Recently, a method for multi-exposure images fusion based on structural decomposition of images into three parts including signal strength, signal structure and signal mean has been introduced. In this paper, we seek to use this decomposition, for images fusion in other fields, including multimodal medical, multi-focus, and infrared and visible images. To increase the fusion quality, besides the introduction of the proposed weighting factor in the structural decomposition, contourlet transformation and the pyramidal structure have also been used. First, each of the K input images are represented into low frequency and high frequency subbands, by using contourlet transform. Then, all the corresponding subbands (resulting from the same scales and directions) are fused with each other, separately and in an iterative process. In this iterative process, first, a separate pyramid structure (including approximation and detail layers) is created for each of the corresponding K subbands. These layers are obtained by the down-sampling of subbands and structural separation based on the proposed new weighting factor. Then, the fusion is performed in the reverse direction of the pyramidal structure and the fused image of the K corresponding subband is obtained. By repeating this process, the fused image will be obtained for all the corresponding subbands. At the end, the final fused image is obtained by the inverse contourlet transformation on the fused images of the subbands. Several visual and quantitative comparisons, with 7 common methods in this field, have been made. In the visual aspect, the proposed method shows the highest quality. In quantitative comparisons based on 6 different criteria, in all three categories of evaluated images (multimodal medical images, multi-focus images, and visible and infrared images), the proposed method ranked first in at least 3 criteria.

© 2023 Published by Semnan University Press.

DOI: <https://doi.org/10.22075/jme.2023.28288.2330>

How to cite this article:

Soleimani, M., Aghagolzadeh, A., & Ezoji, M. (2023). Pyramid Image Fusion Based on Contourlet and Enhanced Structural Decomposition. *Journal of Modeling in Engineering*, 21(73), 71-92. doi: 10.22075/jme.2023.28288.2330

ترکیب هر می تصاویر مبتنی بر تبدیل کانتورلت و تفکیک ساختاری بهبود یافته

مجتبی سلیمانی^۱، علی آقاگل زاده^{۲*} و مهدی ازوجی^۳

چکیده	اطلاعات مقاله
<p>اخیرا، روشی برای ترکیب تصاویر چندنوری بر مبنای تفکیک ساختاری تصاویرها به سه بخش، شامل قدرت سیگنال، ساختار سیگنال و میانگین سیگنال، ارائه شده است. در این مقاله، به دنبال استفاده از این نحوه تفکیک، برای ترکیب تصاویرها در حوزه‌های دیگر، از جمله تصاویر چندحالتی پزشکی، تصاویر چندکانونی و تصاویر مادون قرمز و مرئی هستیم. برای افزایش کیفیت ترکیب، علاوه بر معرفی ضریب وزنی پیشنهادی در تفکیک ساختاری، از تبدیل کانتورلت و تشکیل ساختار هر می نیز استفاده شده است. ابتدا با به کارگیری تبدیل کانتورلت، هر یک از K تصویر اولیه ورودی، به زیرباندهای فرکانس پایین و فرکانس بالا، بازنمایی می‌گردند. سپس، تمامی زیرباندهای متناظر (حاصل از مقیاس‌ها و جهت‌های یکسان در کانتورلت)، به صورت مجزا و در یک روند تکراری، با یکدیگر ترکیب می‌شوند. در این روند تکراری، ابتدا برای هر یک از K زیرباند متناظر، یک ساختار هر می مجزا (شامل لایه‌های تقریب و جزئیات) ایجاد می‌گردد. این لایه‌ها، با نمونه برداری کاهشی زیرباندها و تفکیک ساختاری مبتنی بر ضریب وزنی جدید پیشنهادی، به دست می‌آیند. سپس، عمل ترکیب، در جهت معکوس ساختار هر می، انجام می‌گیرد و تصویر ترکیب شده مربوط به K زیرباند متناظر مدنظر، حاصل می‌گردد. با تکرار این روند، تصویر ترکیب شده برای تمامی زیرباندهای متناظر، به دست خواهد آمد. در انتها، تصویر ترکیب شده نهایی، با اعمال تبدیل کانتورلت معکوس، روی تصویرهای ترکیب شده زیرباندها، حاصل می‌گردد. مقایسه‌های متعدد بصری و کمی با ۷ روش متداول در این حوزه انجام شده است. از لحاظ بصری، روش پیشنهادی بالاترین کیفیت را ارائه می‌دهد. در مقایسه‌های کمی نیز که بر اساس ۶ معیار مختلف صورت گرفته است، در هر سه دسته تصاویر مورد ارزیابی (تصاویر چندحالتی پزشکی، تصاویر چندکانونی و تصاویر مرئی و مادون قرمز)، روش پیشنهادی حداقل در ۳ معیار، در رتبه اول قرار گرفته است.</p>	<p>نوع مقاله: پژوهشی دریافت مقاله: ۱۴۰۱/۰۶/۱۳ بازنگری مقاله: ۱۴۰۱/۱۰/۱۰ پذیرش مقاله: ۱۴۰۱/۱۱/۱۱</p> <p>واژگان کلیدی: ترکیب تصاویر، تفکیک ساختاری تصویر، تبدیل کانتورلت، ساختار هر می، ضریب وزنی پیشنهادی.</p>

۱- مقدمه

که بیشترین اطلاعات از تصویرهای اولیه را در خود داشته باشد. تصویرهای اولیه می‌توانند در قالب‌های مختلف از جمله چندحالتی^۱، چندکانونی^۲، چندحسگری^۳ و چندنوری^۴ باشند. ترکیب تصویرها، کاربردهای مختلفی از جمله در تشخیص پزشکی، عکاسی، شناسایی اشیاء و صنایع نظامی و

امروزه ترکیب تصویرها موضوعی رو به رشد در کاربردهای پردازش تصویر است. ترکیب تصویرها، فرآیندی است که در آن، دو یا چند تصویر دریافتی اولیه از حسگرهای مختلف، با یکدیگر ادغام شده و یک تصویر ترکیبی نهایی می‌سازند

* پست الکترونیک نویسنده مسئول: Aghagol@nit.ac.ir

۱. دانشجوی دکتری دانشگاه صنعتی نوشیروانی بابل

۲. استاد دانشگاه صنعتی نوشیروانی بابل

۳. دانشیار دانشگاه صنعتی نوشیروانی بابل

² Multi-modal

³ Multi-focus

⁴ Multi-sensor

⁵ Multi-exposure

نظارتی دارد. در این مقاله، ترکیب تصویرهای چندحالتی پزشکی، ترکیب تصویرهای چندکانونی و ترکیب تصویرهای مرئی و مادون قرمز مدنظر قرار گرفته است. هر یک از تصویرهای چندحالتی حوزه پزشکی مانند CT^1 ، MRI^2 ، PET^3 و $SPECT^4$ ، اطلاعات متفاوتی در اختیار قرار می‌دهند. CT اطلاعات بافت‌های سخت بدن، MRI بافت‌های نرم بدن، PET عمل کرد یک عضو بدن و $SPECT$ جریان خون را در یک منطقه خاص نشان می‌دهند. در تصویرهای چندکانونی، تصویرهای منبع اولیه، با استفاده از یک دوربین، از یک مکان ثابت و با نقاط کانون متفاوت ثبت می‌شوند. ضعف دوربین‌ها این است که نمی‌توانند به‌طور هم‌زمان بر اشیاء با فواصل مختلف تمرکز کنند و بر این اساس، یا بر اشیاء موجود در پیش‌زمینه و یا بر پس‌زمینه متمرکز هستند. نوع دیگری از کاربردهای تصویرهای چندحالتی، ترکیب تصویرهای مادون قرمز و مرئی است. تصویرهای مرئی، جزئیات تصویری منطبق بر سیستم بینایی انسان را فراهم می‌کنند و دوربین‌های مادون قرمز، اطلاعات تابشی حرارتی را در اختیار قرار می‌دهند. در هر یک از سه دسته تصویرهای قید شده، تصویر ترکیب شده باید بیشترین اطلاعات موجود در تصویرهای اولیه منبع را داشته باشد و در عین حال کمترین نویز و اعوجاج‌های نامطلوب را در بر گیرد [۱].

روش‌های ترکیب تصویرها می‌توانند به دو دسته کلی مکانی و فرکانسی تفکیک گردند. روش‌های حوزه مکان، مستقیماً از مقادیر پیکسل‌های تصویرهای ورودی استفاده کرده و با بهره‌گیری از این پیکسل‌ها، سعی می‌کنند به بهترین خروجی دست یابند. در حوزه فرکانس، تصویرهای اولیه با تبدیل‌هایی به ضرایب چندگانه و چندمقیاس^۵ تبدیل شده و تحلیل‌ها روی آنها صورت می‌گیرند. به‌طور کلی، روش‌های حوزه مکان از فرآیندی ساده و قابل درک برخوردار بوده و از دیگر مزایای آن می‌توان به بار محاسباتی اندک و سرعت در اجرا اشاره نمود [۲]. از جمله معایب این دسته روش‌ها، کاهش تفکیک‌پذیری، تارشدگی، اعوجاج طیفی و

درهم ریختگی رنگ‌ها هستند. نسبت سیگنال به نویز بالاتر، کاهش میزان اعوجاج طیفی، وضوح مکانی بالا و همچنین بهبود نتایج در صورت استفاده از ترکیب چندسطحی، از مزایای روش‌های حوزه فرکانس هستند [۳]. از معایب این روش‌ها نیز می‌توان به این موارد اشاره کرد که: در برخی روش‌ها وضوح تصویر خروجی پایین‌تر از وضوح تصویرهای اولیه است؛ روند پیچیده‌تر و بار محاسباتی بالاتری نسبت به روش‌های مکانی دارند و همچنین کیفیت تصویر خروجی وابستگی زیادی به روش ترکیب انتخاب شده دارد.

برخی روش‌های حوزه مکان عبارتند از: میانگین‌گیری وزن دار [۳]، حداقل مربعات وزن دار [۴]، مشتق جزئی [۵]، K نزدیک‌ترین همسایه (KNN^6) [۶]، استفاده از فیلتر دو طرفه متقاطع^۷ [۷]، شیب^۸ و ریخت‌شناسی^۹ [۸]، فیلتر $gradientlet$ [۹]، شبه فیلتر دو طرفه متقاطع [۱۰]، تحلیل مولفه اصلی^{۱۱} [۱۱] و بازنمایی دو مقیاسی [۱۲]. در [۳]، برای محاسبه وزن‌ها در میانگین‌گیری وزن دار، از بسط $K-L$ استفاده می‌شود. در [۴]، از یک چارچوب بهینه‌سازی مبتنی بر حداقل مربعات وزن دار، برای بهبود نگاشت وزن استفاده شده که در آن ویژگی‌های بافت و اشباع رنگ به کارگرفته می‌شوند. مرجع [۵] نیز از معادلات مشتق جزئی مرتبه چهارم تصویر به همراه تحلیل مولفه اصلی، برای ترکیب تصویرها بهره می‌گیرد. در [۶]، از فیلترهای لاپلاس برای محاسبه برجستگی پیکسل‌ها استفاده شده و سپس از KNN برای یافتن نقشه وزن بهینه هر تصویر، به‌کارگرفته شده است. حساسیت به نویز نقطه ضعف اصلی [۳-۶] است. ترکیب مبتنی بر ارزش پیکسل، با استفاده از فیلتر دو طرفه متقاطع ($PSCBF$) در [۷] انجام شده است. در این روش، با استفاده از فیلتر دو طرفه متقاطع، تصویر جزئیات حاصل می‌گردد و سپس بر اساس قدرت جزئیات، وزن‌ها و ارزش پیکسل‌ها مشخص می‌گردند. در نهایت با میانگین‌گیری وزن دار بر اساس وزن‌های به‌دست آمده، تصویر ترکیب شده نهایی حاصل می‌گردد. مشکلات این روش ایجاد اثر بلوکی در تصویر و بار محاسباتی بسیار بالای آن است. در روش $FFIF^{12}$ ، از اندازه شیب گسسته برای

¹ Computed tomography

² Magnetic resonance imaging

³ Positron emission tomography

⁴ Single-photon emission computed tomography

⁵ Multi-scale

⁶ K nearest neighbor

⁷ Cross bilateral filter

⁸ Gradient

⁹ Morphology

¹⁰ Principal component analysis

¹¹ Karhunen-Loeve expansion

¹² Fast Filtering Image Fusion

حفظ جزئیات و برجستگی‌های تصویر استفاده می‌شود. به منظور حفظ بهتر انرژی تصویر، ترکیب لایه تقریب نیز براساس مشتق انرژی محلی وزن دار صورت می‌گیرد. در پایان نیز با ادغام لایه تقریب و جزئیات، تصویر ترکیب شده نهایی ایجاد می‌گردد. عدم حفظ مناسب لبه‌ها و منحنی‌ها از جمله معایب این روش است.

نمونه‌هایی از روش‌های حوزه فرکانس شامل فیلتر هدایت شده^۴ [۱۳]، تجزیه مقدار منفرد^۵ [۱۴]، همدوسی فاز^۶ [۱۵]، توصیف‌گر SIFT^۷ [۱۶]، تبدیل کانتورلت بدون زیرنمونه‌برداری^۸ [۱۷]، بازنمایی تنک^۹ [۱۸]، تبدیل شیرلت بدون زیرنمونه‌برداری (NSST^{۱۰}) [۱۹]، تبدیل Curvelet [۲۰]، تبدیل موجک مختلط [۲۱]، تبدیل موجک همساز کسینوسی گسسته^{۱۱} [۲۲]، تصویر هدایت شده چندمقیاس^{۱۲} [۲۳] و تفکیک ساختاری چندمقیاس [۲۴] هستند. در [۱۳]، تصویر به دو لایه تقریب و جزئیات تفکیک شده و سپس از فیلتر هدایت شده، برای یافتن میانگین وزن دار به منظور ترکیب این لایه‌ها استفاده می‌شود. عیوب این روش، بار محاسباتی زیاد و پیچیدگی در محاسبه وزن است. تجزیه مقدار منفرد چندوضوحی، برای ترکیب سریع و کارآمد، در [۱۴] به کار رفته است. مرجع [۱۵] روشی برای ترکیب تصویرهای چندکانونی معرفی کرده که از همدوسی فاز با استفاده از تبدیل موجک Gabor مختلط، برای تشخیص میزان تمرکز استفاده می‌کند. بار محاسباتی و عدم حفظ مناسب لبه‌ها، از معایب این روش است. توصیف‌گر SIFT برای استخراج اطلاعات تفکیک‌پذیری محلی و فیلتر هدایت‌شده برای حذف نویز و گسستگی‌های نگاشت وزن، در [۱۶] استفاده شده‌اند. حساسیت به نویز در این روش اثرگذار است. مرجع [۱۷] از تبدیل کانتورلت برای ترکیب تصاویر بهره می‌گیرد. پس از اعمال تبدیل کانتورلت روی تصاویر اولیه، زیرباندهای فرکانس پایین، با استفاده از یک طرح انتخاب مبتنی بر میانگین‌گیری، ترکیب می‌شوند و برای ترکیب زیرباندهای فرکانس بالا، از تفکیک‌پذیری جهت‌دار و انحراف استاندارد

شناسایی تفکیک‌پذیری^۱ و نقاط تیز تصویر استفاده شده است [۸]. سپس با استفاده از یک عملیات فیلتر سریع مبتنی بر ریخت‌شناسی، جاهای خالی و حفره‌ها پر می‌شوند. همچنین در حوزه مکان، از یک فیلتر حفظ‌کننده ساختار برای به دست آوردن نگاشت وزن مطلوب استفاده شده است. حساسیت به نویز و ضعف در تفکیک پس‌زمینه و پیش‌زمینه در این روش دیده می‌شود. در روش GFF^۲ نیز یک تابع آستانه‌گذاری فازی معرفی شده که می‌تواند شیب بافت را از شیب لبه جدا کند [۹]. با استفاده از این آستانه، فیلتری بنام gradientlet معرفی شده است که می‌تواند در عین حال که لبه‌های تصویر و توزیع شدت روشنایی آن را حفظ می‌کند، بافت‌ها با شیب کوچک را حذف نماید. بر اساس فیلتر پیشنهادی و برجستگی تصویر^۳، ترکیب به نحوی انجام می‌شود که از تار شدن جزئیات تصویر در اثر نویز جلوگیری شده و این جزئیات تصویر واضح‌تر شوند. مشکل این روش، عدم حفظ لبه‌ها و تفکیک پس‌زمینه از پیش‌زمینه، در قسمت‌هایی از تصویر است که نوسانات شدت روشنایی زیاد است. در [۱۰] تصویر به دو لایه انرژی و ساختار تبدیل می‌شود. سپس برای ترکیب لایه‌های انرژی، از نگاشت تشخیص برجستگی بصری استفاده می‌شود که اطلاعات جزئیات تصویر را به خوبی حفظ می‌کند. ترکیب لایه‌های ساختار که حاوی جزئیات تصویر هستند نیز با استفاده از روش‌های ریخت‌شناسی چندسطحی و لاپلاس بهبودیافته انجام می‌گیرد. حساسیت به نویز و کاهش کنتراست از معایب این روش است. ترکیب تصاویر با استفاده از تحلیل مولفه اصلی و تغییرات کلی بهینه در [۱۱] استفاده شده است. از یک مدل تغییرات کلی بهینه نرم LI برای قید گذاشتن بین تصویر ترکیب شده و تصویر حاصل از تحلیل مولفه اصلی استفاده شده است. تخریب طیفی و اعوجاج رنگ از جمله معایب روش‌هایی است که بر مبنای تحلیل مولفه اصلی عمل می‌کنند. در [۱۲]، ابتدا تصویر با استفاده از یک تبدیل دو مقیاسی، به دو لایه تقریب و جزئیات تبدیل می‌گردد. سپس برای ترکیب لایه جزئیات، از روشی پیشنهادی برای بهبود و

⁷ Scale-invariant feature transform

⁸ Non-subsampled contourlet transform

⁹ Sparse Representation

¹⁰ Non-subsampled shearlet transform

¹¹ Discrete cosine harmonic wavelet transform

¹² Multi-scale guided image

¹ Contrast

² Gradientlet filter fusion

³ Image saliency

⁴ Guided filter

⁵ Singular value decomposition

⁶ Phase congruency

تصویرهای اولیه پس از اعمال این تبدیل موجک، به زیرباندهای مختلف بازنمایی می‌شوند. در فضای هر یک از زیرباندها، ترکیب وزن دار مبتنی بر انرژی صورت می‌گیرد. زیرباندهای ترکیب شده، به وسیله تبدیل معکوس، تبدیل به تصویر ترکیب شده نهایی می‌گردند. درهم ریختگی طیفی، سایه دار شدن تصویر و بار محاسباتی بالا از معایب این روش است. در MGF، یک بازنمایی هرمی چندمقیاس، با استفاده از فیلتر تصویر هدایت شده ایجاد می‌گردد [۲۳]. سپس نگاشت‌های برجستگی و نگاشت‌های وزنی، با استفاده از لایه‌های جزئیات این هرم، محاسبه می‌گردند. بر اساس نگاشت‌های مذکور، لایه‌های جزئیات با یکدیگر ترکیب شده و در نهایت ساختار هرمی معکوس، منجر به تصویر ترکیب شده نهایی می‌گردد. در هم ریختگی فرکانسی و سایه دار شدن تصویر از عیوب این روش است. مرجع [۲۴]، روش FMSDF⁵ را برای ترکیب تصویرهای چندنوری با صحنه‌های ایستا و پویا، پیشنهاد کرده است. در این روش، ابتدا تصویرهای اولیه با استفاده از روش تفکیک تصویر SPD⁶ [۲۵] که اخیراً برای ترکیب تصویرهای چندنوری معرفی شده است، به سه بخش قدرت سیگنال، ساختار سیگنال و میانگین سیگنال تفکیک می‌گردند. سپس ترکیب تصویرهای، مبتنی بر یک ساختار هرمی صورت می‌گیرد. در عین حال سرعت تفکیک اولیه تصویر نیز در FMSDF بهبود قابل ملاحظه‌ای یافته است.

امروزه روش‌های مبتنی بر یادگیری عمیق نیز با سرعت زیادی در تمامی زمینه‌های پردازش تصویر رو به افزایش هستند [۲۸ - ۲۶]. از جمله این روش‌ها می‌توان به استفاده از CNN⁷ [۲۹] و UDCGAN⁸ [۳۰] برای ترکیب تصویرهای چندکانونی، SAE⁹ [۳۱] و DenseNet¹⁰ [۳۲] برای ترکیب تصویرهای پزشکی و GAN¹¹ برای ترکیب تصویرهای مرئی و مادون قرمز [۳۳]، اشاره کرد. دو مشکل اصلی این روش‌ها، کمبود اطلاعات و تصاویر بسیار زیاد برای آموزش و همچنین نبود حقیقت مبنای¹² است. بازنمایی SPD که در ترکیب تصاویر چندنوری به کار گرفته شده است [۲۵]، با توجه به اینکه اطلاعات شدت روشنایی

بردار جهتی استفاده می‌شوند. در انتها تبدیل کانتورلت معکوس با استفاده از زیرباندهای ترکیب شده صورت می‌گیرد. استفاده از بازنمایی تنک در [۱۸] صورت گرفته است. تصویر به دو لایه تقریب و جزئیات تبدیل شده و از بازنمایی تنک برای ترکیب لایه‌های تقریب و از عملگر فعالیت انرژی همسایه‌ها¹ برای ترکیب لایه جزئیات استفاده شده است. با توجه به اینکه در بازنمایی تنک، تمامی جزئیات ریز و لبه‌ها در برگرفته نمی‌شوند، تصویر ترکیب شده نهایی ممکن است جزئیات بسیار ریز را دارا نباشد. در [۱۹]، با استفاده از تبدیل شیرلت بدون زیرنمونه‌برداری، تصویر تبدیل به لایه‌های فرکانس پایین و فرکانس بالا می‌گردد. برای ادغام لایه‌های فرکانس بالا و پایین، از یک روش مبتنی بر بهینه محلی استفاده شده است. برای ترکیب لایه‌های فرکانس پایین، از فیلتر هم‌تفاقی² و در ترکیب لایه‌های فرکانس بالا، از مجموع لاپلاسین بهبود یافته برای حفظ لبه‌های تصویر استفاده می‌شود. در پایان نیز با استفاده از تبدیل معکوس شیرلت، تصویر ترکیب شده حاصل می‌شود. در [۲۰]، تبدیل Curvelet برای ترکیب تصویرهای چند طیفی به کار گرفته شده است. در این روش (CVT³)، با استفاده از تحلیل چندوضوحی، باندهای چندطیفی که دارای وضوح پایینی هستند، بازنمونه‌برداری می‌شوند تا به تصویرهایی با وضوح بالاتر مطلوب و با لبه‌های واضح تر و تیز شده تبدیل گردند. این فرآیند با افزودن جزئیات بالاگذر جهت دار که با استفاده از تبدیل curvelet به دست می‌آید، انجام می‌گیرد. درهم ریختگی طیفی را می‌توان از جمله معایب این روش دانست. در DTCWT، برای به دست آوردن یک نگاشت ناحیه، از یک تبدیل موجک مختلط دوشاخه‌ای⁴ برای قطعه بندی ویژگی‌های تصویرهای ورودی استفاده می‌شود [۲۱]. پس از محاسبه خصوصیات هر ناحیه، از روشی مبتنی بر ناحیه در حوزه موجک، برای ترکیب تصویرها استفاده شده است. از عیوب این روش، سایه دار شدن تصویر و عدم حفظ تفکیک پذیری در نواحی مختلف تصویر است. در [۲۲]، تبدیل موجک هم ساز گسسته دو بعدی (DCHWT) ارائه شده و

⁸ Generation adversarial network with dense connections

⁹ Attacked autoencoder

¹⁰ Dense Convolutional Network

¹¹ Generative adversarial network

¹² Ground truth

¹ Neighbor energy activity operator

² Co-occurrence filter

³ Curvelet transform

⁴ Dual-tree complex wavelet transform

⁵ Fast multi-scale structural patch decomposition fusion

⁶ Structural patch decomposition

⁷ Convolution neural network

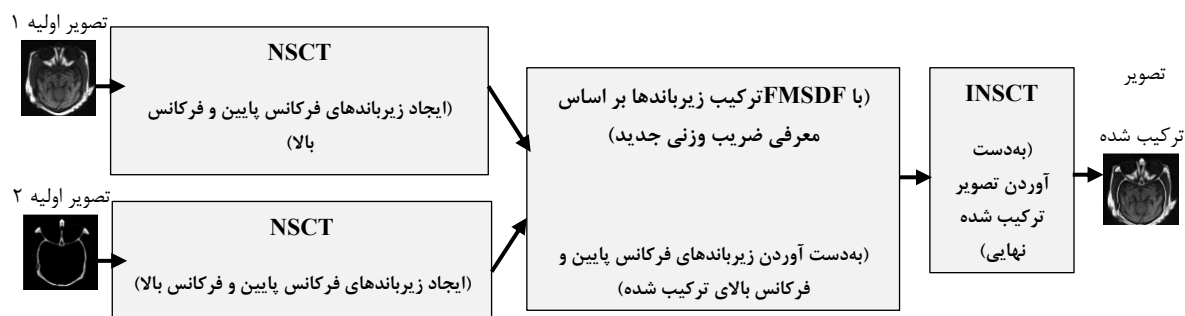
زیرباندهای فرکانس پایین و فرکانس بالا بازنمایی می‌گردند. سپس تمامی جفت زیرباندهای متناظر به صورت مجزا و با استفاده از FMSDF ترکیب می‌گردند. برای بهبود کیفیت ترکیب، ضریب وزنی جدیدی برای استفاده در FMSDF معرفی شده است که جزئیات گوشه‌ها و لبه‌ها را با تاکید بیشتری در نظر می‌گیرد. پس از محاسبه تصویرهای ترکیب شده حاصل از تمامی جفت زیرباندها، تصویر ترکیب شده نهایی با اعمال تبدیل معکوس NSCT² (INSCT) حاصل می‌شود. ارزیابی‌ها در سه دسته مختلف از جمله ترکیب تصویرهای چندحالتی پزشکی، ترکیب تصویرهای چندکانونی و ترکیب تصویرهای مادون قرمز و مرئی انجام شده است. مقایسه با ۷ روش متعارف برای ترکیب تصویرها و بر اساس ۶ معیار کیفی و کمی مختلف، دلالت بر عملکرد مطلوب و برتری روش پیشنهادی در اغلب معیارهای مذکور دارد. در ادامه مقاله، در بخش ۲، پیش‌نیازها و روش‌های مرتبط مورد بررسی قرار می‌گیرند. روش پیشنهادی و ریاضیات آن، در بخش ۳ ارائه می‌گردند. سپس در بخش ۴، حاصل بررسی و مقایسه‌های به عمل آمده با روش‌های متداول و بر اساس چند معیار کمی آورده شده است. در بخش پایانی نیز نتیجه‌گیری کلی بیان می‌شود.

۲- پیش‌نیازها

در این قسمت، تبدیل NSCT [۳۶] و روش FMSDF [۲۴] به عنوان پیش‌نیازهای روش پیشنهادی، بررسی می‌شوند.

۲-۱- تبدیل NSCT

NSCT نوعی بازنمایی چندمقیاسی و چندجهتی برای استفاده در تصویرهای گسسته است [۳۶]. این تبدیل می‌تواند به دو مرحله مختلف تقسیم بندی شود که عبارتند از:

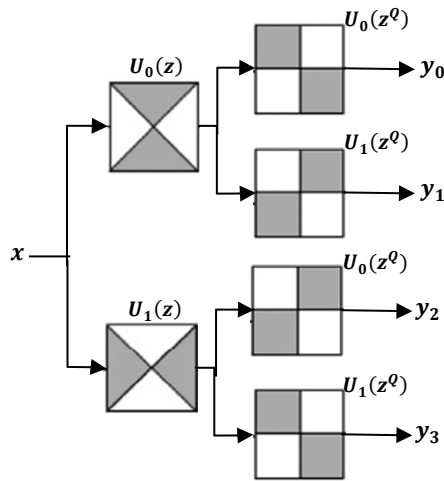


شکل ۱- چارچوب کلی روش پیشنهادی

و ساختار را به طور هم‌زمان حفظ می‌کند و در برابر نویز به خوبی مقاوم است، کیفیت بصری مطلوبی در تصویر نهایی ارائه داده و نتایج کمی آن نیز با روش‌های روزآمد برابری می‌کند. علیرغم مطلوبیت ذکر شده، بار محاسباتی بالا و سایه‌دار شدن تصویر از جمله معایب این روش است. FMSDF در [۲۴]، بار محاسباتی SPD را به شدت کاهش داده و با بهره‌گیری از رویکرد چندمقیاسی و ساختار هرمی، اثر سایه‌دار شدن آن را به خوبی بهبود داده است. بنابراین، FMSDF نسبت به روش‌های بررسی شده قبلی، بسیاری از مشکلات را پوشش داده است. با توجه به مزایای FMSDF و همچنین وجود تشابه نسبی ترکیب تصویرهای چندکانونی با ترکیب تصویرهای چندحالتی و چندکانونی (در آنها کیفیت بصری، شدت روشنایی، تفکیک‌پذیری و حفظ لبه‌ها در تصویر ترکیب شده نهایی مهم است)، در روش پیشنهادی این مقاله، به دنبال استفاده از FMSDF در این حوزه‌های جدید ترکیب تصاویر هستیم. در روش پیشنهادی، از تفکیک SPD و ساختار هرمی با استفاده از یک ضریب وزنی پیشنهادی که تاکید بر جزئیات و لبه‌های تصویر دارد استفاده شده است. جهت بهبود کیفیت خروجی، تبدیل کانتورلت بدون زیرنمونه‌برداری^۱ (NSCT) نیز به کارگیری شده است. با توجه به بهره‌گیری NSCT از مقیاس‌ها و جهت‌های مختلف، این تبدیل می‌تواند ساختارهای مکانی پیچیده را نیز پوشش داده و در نتیجه دو تصویر اولیه ورودی، با استفاده از تبدیل کانتورلت به جهت‌ها، لبه‌ها و جزئیات را به خوبی در نظر بگیرد و بر این اساس، باعث بهبود کیفیت تصویر ترکیب شده نهایی گردد [۳۴] و [۳۵]. چارچوب کلی روش پیشنهادی برای ترکیب دو تصویر (امکان ترکیب هم‌زمان بیش از دو تصویر، K تصویر، نیز در نظر گرفته شده است)، طبق شکل (۱) است.

² Inverse NSCT

¹ Non-subsampled



شکل ۳- بانک فیلتری جهت‌دار بدون زیرنمونه‌برداری چهار کاناله. $U_i(Z), i = 0, 1$ فیلترهای دوبعدی جهت‌دار در حوزه تبدیل z بوده و $U_i(Z^Q), i = 0, 1$ در مرحله دوم، فیلترهای دوبعدی با پوشش صفحه شطرنجی هستند [۳۶].

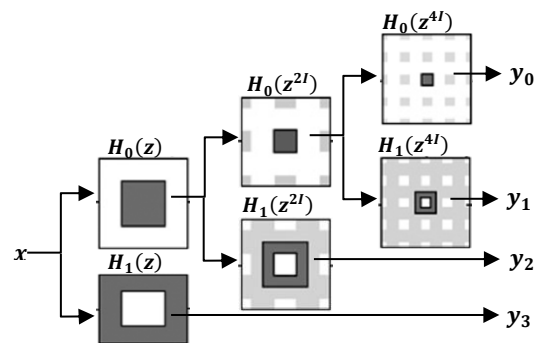
۲-۲- روش FMSDF

در [۲۴]، روش FMSDF برای ترکیب تصویرهای چندنوری^۴ ارائه شده است. تصویرها، با به‌کارگیری و اعمال وصله‌های همپوشان تحلیل می‌گردند. طبق روابط زیر، در بازنمایی^۵ SPD، هر وصله $x \in \mathbb{R}^N$ از تصویر (N تعداد کل پیکسل‌های هر وصله است)، به سه جزء میانگین شدت روشنایی^۶، قدرت سیگنال^۷ و ساختار سیگنال^۸ تفکیک می‌شود:

$$x = l \cdot \mathbf{1} + \|x - l\| \cdot \frac{x - l}{\|x - l\|} = l \cdot \mathbf{1} + \|\tilde{x}\| \cdot \frac{\tilde{x}}{\|\tilde{x}\|} = l \cdot \mathbf{1} + c \cdot s \quad (1)$$

که l یک عدد و میانگین شدت روشنایی وصله x ، $\mathbf{1}$ بردار N بعدی تماماً ۱، \tilde{x} وصله حاصل از حذف میانگین x ، و $\|\cdot\|$ بیانگر نرم l_2 است. $c = \|\tilde{x}\|$ عددی است که نشانگر قدرت سیگنال است. $s = \frac{\tilde{x}}{\|\tilde{x}\|}$ نیز بردار یکه است که جهت آن، به‌نوعی ساختار سیگنال را نشان می‌دهد. در صورتی که K تصویر چندنوری مختلف از یک صحنه در اختیار داشته باشیم، هر یک به‌صورت مجزا وصله بندی شده و اجزاء l, c و s برای هر وصله محاسبه می‌گردد. سپس اجزاء متناظر مربوط به K تصویر اولیه، به‌صورت جداگانه ترکیب

هرم بدون زیرنمونه‌برداری (NSP^۱) و بانک فیلتر جهت‌دار بدون زیرنمونه‌برداری (NSDFB^۲). در مرحله NSP، با استفاده از بانک فیلتر بدون زیرنمونه‌برداری دو کاناله، ویژگی چندمقیاسه ایجاد می‌گردد. در هر مرحله تفکیک NSP، یک تصویر فرکانس پایین و یک تصویر فرکانس بالا تشکیل می‌شوند. مرحله تفکیک NSP با تعداد مراحل مدنظر و به‌صورت پشت‌سرهم صورت می‌گیرد. در صورتی که NSP شامل P مرحله ($p = 1, 2, \dots, P$) مختلف تفکیک باشد، در مجموع $P + 1$ زیرتصویر هم‌اندازه با تصویر اصلی، شامل یک تصویر فرکانس پایین و P تصویر فرکانس بالا، ایجاد می‌گردند. شکل (۲) تصویر نمونه NSP با $P = 3$ مرحله را نشان می‌دهد. NSDFB هم‌بانک‌های فیلتری بدون زیرنمونه‌برداری دو کاناله هستند که با ترکیب بانک‌های فیلتری بادبزی^۳ جهت‌دار ساخته می‌شوند. NSDFB امکان تفکیک جهت‌دار D مرحله‌ای ($d = 1, 2, \dots, D$) را روی تصویرهای فرکانس بالای خروجی NSP، در هرکدام از مقیاس‌ها، فراهم می‌کند. بنابراین، 2^D زیرتصویر جهت‌دار که هم‌اندازه تصویر اصلی هستند ایجاد می‌کند. در نتیجه، NSDFB به‌نوعی ویژگی چندجهته بودن را برای NSCT برآورده می‌کند که منجر به کسب اطلاعات جزئیات جهتی دقیق‌تری می‌گردد. شکل (۳) نمونه‌ای از یک NSDFB چهار کاناله است که با استفاده از بانک‌های فیلتری بادبزی دو کاناله ساخته شده‌اند.



شکل ۲- تفکیک هرمی بدون زیرنمونه‌برداری سه مرحله‌ای. $H_i(Z), i = 0, 1$ فیلترهای دوبعدی پایین‌گذر و بالاگذر و $H_i(Z^{2^i}), i = 0, 1$ و $H_i(Z^{2^i}), i = 0, 1$ به‌ترتیب فیلترهای دوبعدی در مرحله دوم و سوم هستند. تمامی فیلترها در حوزه تبدیل z هستند [۳۶].

^۵ Structural patch decomposition

^۶ Mean intensity

^۷ Signal strength

^۸ Signal structure

^۱ Non-sampled pyramid

^۲ Non-sampled directional filter bank

^۳ Fan filter

^۴ Multi Exposure

اصلاحیه رابطه β_k : با توجه به رابطه (۶) و (۷) که در مرجع [۲۴] قید شده است، شرط $\sum_k \beta_k = 1$ نقض می‌گردد و رابطه (۷) درست نیست. بر اساس روابط مطرح شده در معرفی SPD [۲۵]، محاسبه $\hat{\mathbf{s}}$ طبق روابط زیر انجام می‌گیرد:

$$\hat{\mathbf{s}} = \frac{\bar{\mathbf{s}}}{\|\bar{\mathbf{s}}\|} \quad \text{و} \quad \bar{\mathbf{s}} = \frac{\sum_{k=1}^K \psi_k s_k}{\sum_{k=1}^K \psi_k} \quad (۸)$$

در این رابطه، $\psi_k = \|\tilde{\mathbf{x}}_k\|^\eta$ که $\eta \geq 0$ است. در [۲۴] اثبات شده که $\|\bar{\mathbf{s}}\| \cong 1$ بوده و در نتیجه $\hat{\mathbf{s}} \approx \bar{\mathbf{s}}$ است. بنابراین:

$$\hat{\mathbf{s}} = \bar{\mathbf{s}} = \frac{\psi_1}{\sum_{k=1}^K \psi_k} s_1 + \frac{\psi_2}{\sum_{k=1}^K \psi_k} s_2 + \dots + \frac{\psi_k}{\sum_{k=1}^K \psi_k} s_k = \beta_1 s_1 + \beta_2 s_2 + \dots + \beta_k s_k \quad (۹)$$

برای درست بودن رابطه (۶) و نقض نشدن شرط $\sum_k \beta_k = 1$ ، رابطه (۷) باید به صورت زیر اصلاح گردد:

$$\beta_k = \frac{\psi_k}{\sum_{k=1}^K \psi_k} = \frac{\|\tilde{\mathbf{x}}_k\|^\eta}{\sum_{k=1}^K \|\tilde{\mathbf{x}}_k\|^\eta}; \quad \eta \geq 0, \quad k = 1, \dots, K \quad (۱۰)$$

با قرار دادن $\hat{\mathbf{L}}$ و $\hat{\mathbf{s}}$ در رابطه (۲)، خواهیم داشت:

$$\hat{\mathbf{x}} \approx \sum_{k=1}^K (a_k l_k \cdot \mathbf{1} + \hat{\mathbf{c}} \cdot \beta_k s_k) = \sum_{k=1}^K (a_k l_k \cdot \mathbf{1} + \hat{\mathbf{c}} \cdot \beta_k \frac{\tilde{\mathbf{x}}_k}{\|\tilde{\mathbf{x}}_k\|}) = \sum_{k=1}^K (a_k l_k \cdot \mathbf{1} + \gamma_k \cdot (\mathbf{x}_k - l_k)) \quad (۱۱)$$

که $\gamma_k = \frac{\hat{\mathbf{c}} \beta_k}{\|\tilde{\mathbf{x}}_k\|}$ است. رابطه بالا را می‌توان، طبق روش زیر نیز محاسبه نمود. در این حالت، محاسبات مستقیماً روی تصویرهای \mathbf{X}_k انجام می‌گیرد و تصویر نهایی حاصل از ترکیب ($\hat{\mathbf{X}}$) طبق رابطه زیر محاسبه می‌گردد:

$$\hat{\mathbf{X}} = \sum_{k=1}^K (\mathbf{L}(a_k \odot \mathbf{L}(\mathbf{X}_k)) + \mathbf{L}(\gamma_k) \odot (\mathbf{X}_k - \mathbf{L}(\gamma_k \odot \mathbf{L}(\mathbf{X}_k)))) \quad (۱۲)$$

در این رابطه، $\mathbf{L}(\cdot)$ فیلتر میانگین با اندازه پنجره N و \odot بیانگر ضرب هادامارد است. فرآیند اعمال فیلتر میانگین، با استفاده از box filter [۳۷] صورت می‌گیرد که بار محاسباتی خطی دارد و باعث کاهش زمان پیاده‌سازی و پیچیدگی محاسباتی می‌گردد. تا اینجا، روش ترکیب سریع مبتنی بر تفکیک ساختاری تصویرها به سه بخش، میانگین شدت روشنایی، قدرت سیگنال و ساختار سیگنال

می‌گردند (محاسبه مجزای $\hat{\mathbf{L}}$ ، $\hat{\mathbf{c}}$ و $\hat{\mathbf{s}}$). در انتها، تصویر ترکیب شده نهایی ($\hat{\mathbf{x}}$)، با استفاده از عمل معکوس تفکیک ساختاری و طبق رابطه زیر حاصل می‌گردد.

$$\hat{\mathbf{x}} = \hat{\mathbf{L}} \cdot \mathbf{1} + \hat{\mathbf{c}} \cdot \hat{\mathbf{s}} \quad (۲)$$

در این رابطه، میانگین شدت روشنایی محلی برای به دست آوردن تصویر ترکیب شده نهایی برابر است با:

$$\hat{\mathbf{L}} = \sum_{k=1}^K a_k l_k; \quad (۳)$$

$$a_k \geq 0 \quad \text{for } k = 1, \dots, K \quad \text{و} \quad \sum_k a_k = 1$$

وزن‌های a_k ، میزان نور مناسب تصویر را نشان داده و طبق رابطه زیر محاسبه می‌گردند.

$$a_k = \frac{\arctan(0.5\lambda - |0.5 - \mathbf{X}_k| \lambda)}{\sum_{k=1}^K \arctan(0.5\lambda - |0.5 - \mathbf{X}_k| \lambda)} \quad (۴)$$

λ پارامتر ثابت و \mathbf{X}_k نیز k امین تصویر است. $\hat{\mathbf{c}}$ قدرت سیگنال محلی مورد نیاز برای به دست آوردن تصویر ترکیب شده نهایی است. شدت روشنایی \mathbf{X}_k باید نرمالیزه شده و سپس در این رابطه مورد استفاده قرار گیرند.

با توجه به اینکه هرچه تصویر شرایط نوری و کیفیت بصری بهتری داشته باشد، تفکیک‌پذیری بالاتری خواهد داشت و تفکیک‌پذیری، رابطه مستقیم با قدرت سیگنال داد، قدرت سیگنال مطلوب، طبق رابطه زیر محاسبه می‌گردد:

$$\hat{\mathbf{c}} = \max_{1 \leq k \leq K} \|\tilde{\mathbf{x}}_k\| = \max_{1 \leq k \leq K} c_k \quad (۵)$$

ساختار سیگنال محلی مطلوب $\hat{\mathbf{s}}$ نیز طبق رابطه زیر محاسبه می‌گردد:

$$\hat{\mathbf{s}} = \sum_{k=1}^K \beta_k s_k; \quad (۶)$$

$$\beta_k \geq 0 \quad \text{for } k = 1, \dots, K \quad \text{و} \quad \sum_k \beta_k = 1$$

وزن‌های β_k بیانگر میزان مشارکت هر یک از وصله‌های تصویرهای اولیه در ساختار وصله تصویر ترکیب شده است. این وزن‌ها، نیز متناسب با قدرت سیگنال و طبق رابطه زیر حاصل می‌گردند.

$$\beta_k = \|\tilde{\mathbf{x}}_k\|^\eta; \quad \eta \geq 0, \quad k = 1, \dots, K \quad (۷)$$

که η یک عدد ثابت دلخواه است و هرچه عددی بزرگتر باشد، تاکید بیشتری بر وصله‌هایی با قدرت بزرگتر خواهد داشت.

H و W به ترتیب بیان گر ارتفاع و عرض تصویرهای اصلی ورودی هستند. تصویر لایه تقریب مدنظر نیز طبق رابطه زیر حاصل می گردد:

$$\hat{B}^{(j)} = \sum_{k=1}^K \mathbf{L}(a_k^{(j)} \odot \mathbf{L}(X_k^{(j)})) \quad (18)$$

تصویر ترکیب شده نهایی، در یک روند تکراری و با نمونه برداری افزایشی و افزودن لایه جزئیات به لایه های میانی تقریب، طبق رابطه زیر حاصل می گردد:

$$\hat{B}^{(j)} = \mathbf{L}(U(\hat{B}^{(j+1)} + \hat{H}^{(j+1)})) ; \quad (19)$$

$$j = 1, 2, \dots, J - 1$$

و در انتها تصویر ترکیب شده نهایی برابر است با:

$$\hat{X} = \hat{B}^{(1)} + \hat{H}^{(1)} \quad (20)$$

۳- روش پیشنهادی (NSCT-SPDF)

روش پیشنهادی که ترکیبی از FMSDF و NSCT است (NSCT-SPDF)²، با هدف ترکیب تصویرهای چندحالتی و چندکانونی ارائه شده است. همانند ترکیب تصویرهای چندنوری، در ترکیب تصویرهای چندحالتی (مانند تصویرهای پزشکی و تصویرهای مرئی و مادون قرمز) و ترکیب تصویرهای چندکانونی نیز، به دنبال ادغام چند تصویر با خصوصیات شدت روشنایی و تفکیک پذیری³ مختلف هستیم، بنابراین می توان از بازنمایی SPD، برای ترکیب این نوع از تصویرها نیز بهره برد. NSCT نیز از مقیاس ها و جهت های مختلف بهره می برد و می تواند ساختارهای مکانی پیچیده را پوشش داده و جهت ها، لبه ها و جزئیات را به خوبی در نظر بگیرد. به همین دلیل در روش پیشنهادی از NSCT استفاده می شود تا با در نظر گرفتن جزئیات بیشتر، باعث بهبود کیفیت تصویر ترکیب شده نهایی گردد [۳۴ و ۳۵].

در روش پیشنهادی، هر یک از K تصویر اصلی ورودی، با استفاده از NSCT، تبدیل به زیرمجموعه ای از تصویرهای، شامل یک تصویر فرکانس پایین X_p و تصویرهای فرکانس بالای $X_{p,d}$ می گردند و برای هر یک از K تصویر مذکور،

(F-SPDF¹) ارائه شده است.

در صورتی که برای هر پیکسل که در مرکز وصله قرار دارد، میانگین شدت روشنایی به طور مجزا محاسبه گردد، رابطه (۱۲) می تواند به شکل زیر تبدیل گردد.

$$\hat{X} = \sum_{k=1}^K (\mathbf{L}(a_k) \odot \mathbf{L}(X_k) + \mathbf{L}(\gamma_k) \odot (\mathbf{X}_k - \mathbf{L}(X_k))) = \hat{B} + \hat{H} \quad (13)$$

که معادل تفکیک تصویر به دو لایه تقریب (\hat{B}) و جزئیات (\hat{H}) است. $\mathbf{L}(a_k)$ و $\mathbf{L}(\gamma_k)$ به ترتیب وزن های نگاشت لایه تقریب و لایه جزئیات هستند.

در [۲۴]، علاوه بر کاهش بار محاسباتی، برای بهبود عملکرد ترکیب نیز یک روش چندمقیاسی با استفاده از SPD ارائه شده که مبنای آن به کارگیری ساختار چندمقیاسی برای تصویر و به کارگیری رابطه (۱۳) برای تحلیل تصویرهای لایه های تقریب و جزئیات است. طبق روابط (۱۲) و (۱۳)، برای هر یک از تصویرهای اصلی اولیه، تصویر جزئیات طبق رابطه زیر حاصل می گردد که این تصویر جزئیات را، تصویر مقیاس اول در نظر می گیریم:

$$\hat{H}^{(1)} = \sum_{k=1}^K (\mathbf{L}(\gamma_k^{(1)}) \odot \mathbf{X}_k^{(1)} - \mathbf{L}(\gamma_k^{(1)} \odot \mathbf{L}(X_k^{(1)}))) \quad (14)$$

برای تشکیل تصویر جزئیات در مقیاس ۲، ابتدا طبق رابطه زیر، عمل نمونه برداری کاهشی روی $\mathbf{L}(X_k^{(1)})$ انجام می گیرد که در آن $D(\cdot)$ بیان گر نمونه برداری کاهشی با فاکتور ۲ است:

$$\mathbf{X}_k^{(2)} = D(\mathbf{L}(X_k^{(1)})) \quad (15)$$

سپس تصویر لایه جزئیات در مقیاس دوم طبق رابطه زیر محاسبه می گردد:

$$\hat{H}^{(2)} = \sum_{k=1}^K (\mathbf{L}(\gamma_k^{(2)}) \odot \mathbf{X}_k^{(2)} - \mathbf{L}(\gamma_k^{(2)} \odot \mathbf{L}(X_k^{(2)}))) \quad (16)$$

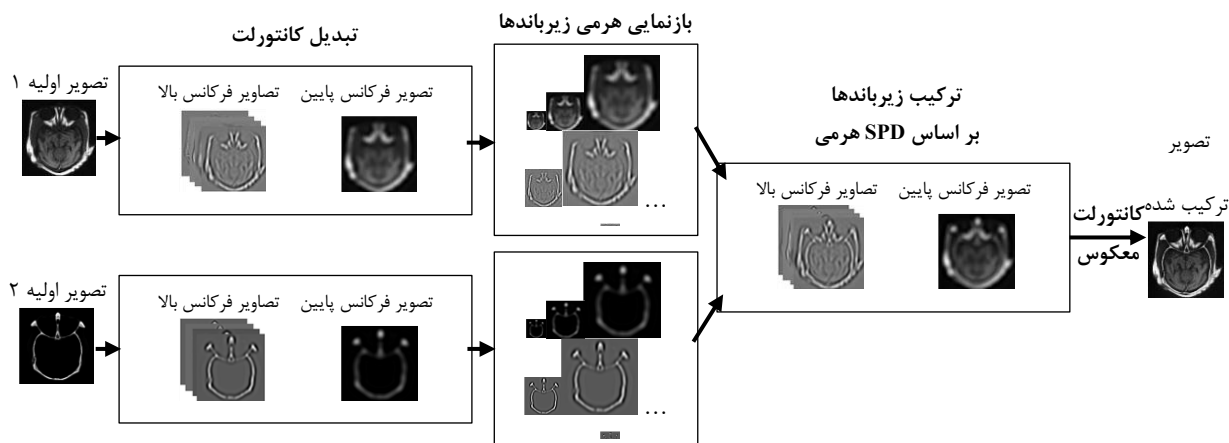
این عمل تکراری، برای محاسبه $X_k^{(j)}$ ها تا رسیدن به درشت ترین مقیاس (J) ادامه می یابد که برابر است با:

$$J = \lfloor \log_2 \min(H, W) \rfloor - 3 \quad (17)$$

³ Contrast

¹ Fast structural patch decomposition fusion

² NSCT based structural patch decomposition fusion



شکل ۴- روندنمای روش پیشنهادی (NSCT-SPDF)

داد:

$$\hat{X}_{p,d} = \sum_{k=1}^K \left(L(a_{p,d,k} \odot L(X_{p,d,k})) + L(\gamma_{p,d,k}) \odot X_{p,d,k} - L(\gamma_{p,d,k} \odot L(X_{p,d,k})) \right) = \hat{B}_{p,d} + \hat{H}_{p,d} \quad (21)$$

$; p = 1, 2, \dots, P, \quad d = 1, 2, \dots, D$



(الف)



(پ)

(ب)

شکل ۵- نمونه‌ای از بهبود ترکیب تصاویر با استفاده از ساختار هرمی (الف) تصاویر اولیه (ب) تصویر ترکیب شده بدون ساختار هرمی (SPD) (پ) تصویر ترکیب شده با استفاده از ساختار هرمی (SPD مبتنی بر ساختار هرمی)

مجموعه $X_k: \{X_{p,k}, X_{p,d,k}\}; k = 1, 2, \dots, K$ حاصل می‌شود. سپس تمامی زیرباندیهای متناظر از این مجموعه، به صورت مجزا و با استفاده از روش FMSDF ترکیب می‌گردند. در واقع، ابتدا K تصویر فرکانس پایین $X_{p,k}; k = 1, 2, \dots, K$ حاصل از K تصویر اصلی ورودی، به صورت مجزا بر اساس FMSDF، ترکیب می‌شوند و تصویر ترکیبی \hat{X}_p به دست می‌آید. سپس K تصویر متناظر فرکانس بالا $X_{2,1,k}; k = 1, 2, \dots, K$ ترکیب می‌گردند (تصویر ترکیبی $\hat{X}_{2,1}$ حاصل می‌شود) و این ترکیب مجزای تصویرهای متناظر، در یک روند تکراری تا به انتها برای تمامی سطوح و جهت‌های p و d صورت می‌گیرد. در پایان، با در اختیار داشتن تصویرهای ترکیب شده فرکانس پایین و فرکانس بالا $\{\hat{X}_p, \hat{X}_{p,d}\}$ ، عمل معکوس NSCT (INSCT¹) انجام شده و تصویر ترکیب شده نهایی \hat{X} حاصل می‌گردد. روندنمای روش پیشنهادی، برای ترکیب ۲ تصویر اولیه، طبق شکل (۴) است. در ادامه، جزئیات ترکیب پیشنهادی برای هر یک از زیرباندیهای متناظر، مبتنی بر ساختار هرمی FMSDF ارائه شده است.

۱-۳- ترکیب زیرباندیها مبتنی بر FMSDF

با توجه به اینکه نحوه ترکیب زیرباندیهای فرکانس پایین و فرکانس بالا، مشابه است، برای سادگی، زیرباندیهای فرکانس پایین به صورت $X_{p,1}$ در نظر گرفته می‌شوند و بنابراین در ادامه $X_{p,d}$ اشاره به تمامی زیرباندیهای فرکانس بالا و پایین دارد. با توجه به رابطه‌های (۱۲) و (۱۳)، تصویرهای حاصل از ترکیب زیرباندیهای متناظر را می‌توان به صورت زیر نمایش

¹ Inverse NSCT

تصویر ترکیب شده نهایی با استفاده از یک فرآیند تکراری حاصل می‌شود. ابتدا مجموع تصویرهای تقریب و جزئیات آخرین مقیاس هرم (J) محاسبه می‌گردد. سپس با نمونه برداری افزایشی ($U(\cdot)$) و اعمال فیلتر میانگین، خروجی حاصل از ترکیب تصویر تقریب، در مقیاس $J - 1$ ایجاد می‌گردد. روند مذکور طبق رابطه (۲۵) ادامه می‌یابد تا اینکه تصویر تقریب مقیاس ۱ حاصل گردد:

$$\hat{B}_{p,d}^{(j)} = L \left(U \left(\hat{B}_{p,d}^{(j+1)} + \hat{H}_{p,d}^{(j+1)} \right) \right); \quad (25)$$

$$j = 1, 2, \dots, J - 1, \quad p = 1, 2, \dots, P, \quad d = 1, 2, \dots, D$$

در انتها تصویر ترکیب زیرباند‌های متناظر مختلف برابر است با:

$$\hat{X}_{p,d} = \hat{B}_{p,d}^{(1)} + \hat{H}_{p,d}^{(1)}; \quad (26)$$

$$p = 1, 2, \dots, P, \quad d = 1, 2, \dots, D$$

حال که تصویرهای ترکیب شده زیرباند‌های متناظر محاسبه شده و مجموعه $\{\hat{X}_p, \hat{X}_{p,d}\}; p = 1, 2, \dots, P, d = 1, 2, \dots, D$ به دست آمده است، عمل معکوس NSCT روی این مجموعه انجام شده و تصویر ترکیب شده نهایی حاصل می‌گردد.

$$INSCT\{\hat{X}_p, \hat{X}_{p,d}\} \rightarrow \hat{X}; \quad (27)$$

$$p = 1, 2, \dots, P, \quad d = 1, 2, \dots, D$$

۲-۳- ضریب وزنی پیشنهادی

برای محاسبه ضریب $a_{p,d_k}^{(j)}$ جدید پیشنهادی، از ترکیب مقطع^۱ گاوسی دو بعدی [۲۵] و معیار $WSEML^2$ [۳۸] استفاده می‌شود.

$$a_{p,d_k}^{(j)} = \frac{\tau_{p,d_k}}{\sum_{k=1}^K \tau_{p,d_k}}; \quad (28)$$

$$k = 1, 2, \dots, K, \quad p = 1, 2, \dots, P, \quad d = 1, 2, \dots, D$$

که τ_{p,d_k} طبق رابطه زیر محاسبه می‌گردد:

$$\tau_{p,d_k} = \exp \left(- \frac{(\mu_{p,d_k} - 0.5)^2}{2\sigma_g^2} - \frac{(\chi_{p,d_k}^{(j)} - 0.5)^2}{2\sigma_l^2} \right) \times WSEML_{p,d_k}; \quad (29)$$

$$k = 1, 2, \dots, K, \quad p = 1, 2, \dots, P, \quad d = 1, 2, \dots, D$$

در آزمایش‌های انجام شده مشاهده گردید که، اگر ترکیب تصویرها صرفاً با استفاده از X_{p,d_k} ها صورت گیرد، تصویر ترکیب شده نهایی کیفیت مناسبی نداشته و هاله‌هایی در آن مشاهده می‌شود. برای بهبود این مساله، از SPD مبتنی بر ساختار هرمی بهره گرفته می‌شود. شکل (۵) وجود هاله‌ها و بهبود آنها با بهره‌گیری ساختار هرمی را نشان می‌دهد. در SPD مبتنی بر ساختار هرمی، ابتدا ساختار هرمی هر یک از X_{p,d_k} ها با استفاده از میانگین‌گیری و نمونه‌برداری کاهشی با ضریب ۲، تشکیل می‌شود. اولین تصویر در ساختار هرمی (در مقیاس ۱) همان تصویر اصلی $X_{p,d_k}^{(1)}$ است و تصویرهای مقیاس‌های بعدی هرم طبق رابطه زیر حاصل می‌گردند:

$$X_{p,d_k}^{(j+1)} = D \left(L \left(X_{p,d_k}^{(j)} \right) \right); \quad (22)$$

$$k = 1, 2, \dots, K, \quad j = 1, 2, \dots, J - 1, \quad p = 1, 2, \dots, P, \quad d = 1, 2, \dots, D$$

L بیان‌گر فیلتر میانگین و $D(\cdot)$ بیان‌گر نمونه‌برداری کاهشی با فاکتور ۲ است. آخرین مقیاس هرم مذکور برابر J و طبق رابطه (۱۷) به دست می‌آید.

پس از تشکیل ساختار هرمی برای تمامی تصویرهای X_{p,d_k} مربوط به K تصویر اولیه ورودی، ترکیب این تصویرهای مبتنی بر ساختار هرمی صورت می‌گیرد. نحوه ترکیب مجزای لایه‌های تقریب و جزئیات مورد نیاز، بر اساس رابطه (۲۱) است. برای ترکیب، تنها لایه تقریب مورد نیاز، لایه تقریب نهایی (مقیاس J) است و ترکیب آن طبق رابطه زیر حاصل می‌گردد:

$$\hat{B}_{p,d}^{(j)} = \sum_{k=1}^K L \left(a_{p,d_k}^{(j)} \odot L \left(X_{p,d_k}^{(j)} \right) \right); \quad (23)$$

$$p = 1, 2, \dots, P, \quad d = 1, 2, \dots, D$$

ترکیب لایه جزئیات باید برای تمامی مقیاس‌های هرم انجام شده و طبق رابطه زیر محاسبه شود:

$$\hat{H}_{p,d}^{(j)} = \sum_{k=1}^K \left(L \left(\gamma_{p,d_k}^{(j)} \right) \odot X_{p,d_k}^{(j)} - L \left(\gamma_{p,d_k}^{(j)} \odot L \left(X_{p,d_k}^{(j)} \right) \right) \right); \quad (24)$$

$$j = 1, 2, \dots, J, \quad p = 1, 2, \dots, P, \quad d = 1, 2, \dots, D$$

² Weighted sum of eight-neighborhood-based modified Laplacian

¹ Profile

و DCHWT [۲۲] مقایسه می‌گردد. کدهای شبیه‌سازی روش‌های مورد مقایسه، در اصل مقالات مذکور و فضای سایبری در دسترس هستند. بررسی‌ها بر اساس پارامترهای اصلی این کدها صورت گرفته‌اند. تمامی شبیه‌سازی‌ها با استفاده از نرم‌افزار MATLAB 2014a انجام شده است.

۴-۱- معیارهای ارزیابی

از معیارهای متداول زیر جهت ارزیابی روش پیشنهادی استفاده شده است:

- Q_G : معیاری مبتنی بر شیب است که میزان انتقال اطلاعات لبه‌ها از تصویرهای اولیه به تصویر ترکیب شده را ارزیابی می‌کند [۳۹]. هر اندازه عدد مذکور بزرگتر باشد، بیان‌گر تصویر ترکیب شده با کیفیت بالاتر است.

- Q_{SF} : معیاری مبتنی بر فرکانس مکانی است که بیان‌گر میزان خطای شیب در جهت‌های مختلف است [۴۰]. این معیار می‌تواند مقادیر مثبت یا منفی داشته باشد و هرچه به عدد صفر نزدیک‌تر باشد، نشان‌دهنده ترکیب بهتر تصویرها است.

- Q_p : در این معیار، با استفاده از همدوسی فاز، مقدار قدرمطلق ویژگی‌های تصویر حاصل می‌گردد [۴۱]. هر اندازه ترکیب تصویرها بهتر انجام گیرد، حاصل این معیار عددی بزرگ‌تر خواهد بود.

- Q_S : کیفیت ترکیب را بر اساس برجستگی و واریانس تصویر محاسبه می‌کند و بزرگ‌تر بودن آن معادل کیفیت بالاتر ترکیب انجام شده است [۴۲].

- Q_{CV} : معیاری الهام گرفته از ادراک انسان است و هر اندازه مقدار کوچک‌تری داشته باشد مطلوب‌تر و معادل ترکیب با کیفیت بالاتر است [۴۳].

- SF : این معیار، سطح فعالیت کلی در تصویر را با محاسبه شیب‌ها، به‌دست می‌آورد [۴۴]. ترکیب مطلوب‌تر منجر به SF بزرگ‌تر خواهد شد.

۴-۲- پایگاه داده و تنظیمات پارامترها

مقایسه و ارزیابی‌ها روی دامنه متنوعی از تصویرهای ورودی صورت گرفته‌اند که عبارتند از:

- تصویرهای چندحالتی پزشکی: تصویرهای مذکور از پایگاه داده‌های Brainweb [۴۵] و Harvard [۴۶] به‌دست آمده‌اند. در مجموع ۱۰ جفت تصویر که حالت‌های مختلف تصویربرداری از جمله CT، MRI،

در این رابطه، μ_{p,d_k} میانگین کل شدت روشنایی $\mathbf{X}_{p,d_k}^{(J)}$ است ($\mathbf{X}_{p,d_k}^{(J)}$ باید در بازه $[0, 1]$ نرمالیزه شده باشد)، σ_g و σ_l گسترش مقطع گاوسی را در راستای ابعاد μ_{p,d_k} و $\mathbf{X}_{p,d_k}^{(J)}$ کنترل می‌کنند. WSEML نیز لاپلاس بهبودیافته با استفاده از ۸ همسایه حول هر پیکسل است و به‌منظور بهره‌گیری از جزئیات گوشه‌ها و لبه‌های تصویر، استفاده شده است. در صورتی که پنجره وزنی \mathbf{W} به‌صورت زیر باشد:

$$\mathbf{W} = \frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix} \quad (30)$$

در $WSEML_{p,d_k}$ پیکسل (i, j) برابر است با:

$$WSEML_{p,d_k}(i, j) = \sum_{u=-1}^1 \sum_{v=-1}^1 \mathbf{W}(u+2, v+2) \times EML_{p,d_k}(i+u, j+v); \quad (31)$$

$$k = 1, 2, \dots, K, \quad p = 1, 2, \dots, P,$$

$$d = 1, 2, \dots, D$$

که EML همان لاپلاسیان بهبودیافته است که از محاسبه شیب ۸ همسایه مجاور و طبق رابطه زیر محاسبه می‌شود [۳۸]. ضریب $\frac{1}{\sqrt{2}}$ برای شیب‌های قطری، بر مبنای فاصله اقلیدسی آنها است.

$$EML_{p,d_k}(i, j) = \left| 2\mathbf{X}_{p,d_k}^{(J)}(i, j) - \mathbf{X}_{p,d_k}^{(J)}(i-1, j) - \mathbf{X}_{p,d_k}^{(J)}(i+1, j) \right| + \left| 2\mathbf{X}_{p,d_k}^{(J)}(i, j) - \mathbf{X}_{p,d_k}^{(J)}(i, j-1) - \mathbf{X}_{p,d_k}^{(J)}(i, j+1) \right| + \frac{1}{\sqrt{2}} \left| 2\mathbf{X}_{p,d_k}^{(J)}(i, j) - \mathbf{X}_{p,d_k}^{(J)}(i-1, j-1) - \mathbf{X}_{p,d_k}^{(J)}(i+1, j+1) \right| + \frac{1}{\sqrt{2}} \left| 2\mathbf{X}_{p,d_k}^{(J)}(i, j) - \mathbf{X}_{p,d_k}^{(J)}(i-1, j+1) - \mathbf{X}_{p,d_k}^{(J)}(i+1, j-1) \right| ;$$

$$k = 1, 2, \dots, K, \quad p = 1, 2, \dots, P,$$

$$d = 1, 2, \dots, D$$

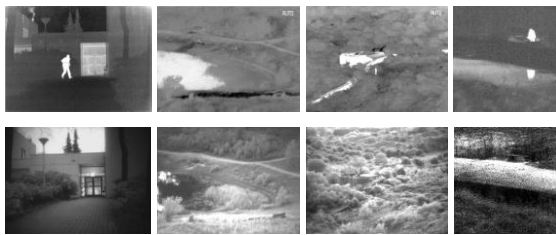
طبق روابط (۲۷-۳۲)، با توجه به اینکه $\tau_{p,d_k} \geq 0$ است، شرط‌های $\sum_{k=1}^K a_{p,d_k}^{(J)} = 1$ و $a_{p,d_k}^{(J)} \geq 0$ برقرار هستند.

۴- بررسی‌ها و نتایج

در این بخش، روش پیشنهادی (NSCT-SPDF) با روش‌های متداول دیگر شامل: CVT [۲۰]، DTCWT [۲۱]، GFF [۹]، MGF [۲۳]، FFIF [۸]، PSCBF [۷]



(پ) تصویرهای چندکانونی



(ت) تصاویر پایگاه داده TNO (سطر اول مادون قرمز و سطر دوم مرئی)



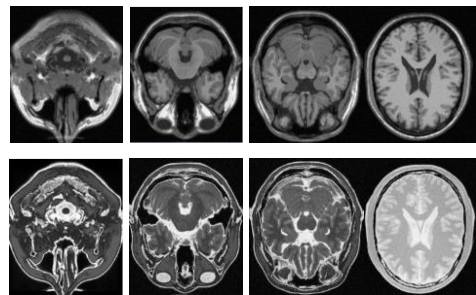
(ث) تصویر مرئی و MMW

شکل ۶- مجموعه ۳۰ تصویر استفاده شده در بررسی‌ها در مجموع ۳۰ جفت تصویر مختلف که همگی آنها به‌طور کامل بر یکدیگر منطبق هستند، برای ارزیابی‌ها به‌کار گرفته شده است. روش پیشنهادی می‌تواند برای ترکیب هم‌زمان بیش از ۲ تصویر ($K > 2$) نیز استفاده گردد، ولی در مقایسه‌های انجام شده، تنها ترکیب ۲ تصویر ($K = 2$) مورد نظر قرار گرفته است. مجموعه تصویرهای به‌کار رفته، در شکل (۶) نشان داده شده‌اند. بر اساس دو مرجع [۳۴] و [۳۵]، مقادیر به این صورت انتخاب شده‌اند که: $\eta = 0.2$ و $\sigma_g = 0.5$ ، $\sigma_l = 0.5$ و چون تصویرهای ورودی خاکستری و دو بعدی هستند، اندازه پنجره به‌کار رفته در تمامی مقیاس‌های هرمی $N = 8 \times 8$ است. بهترین خروجی‌ها در تبدیل NSCT، به ازای $P = 4$ و $D = 1$ به‌دست

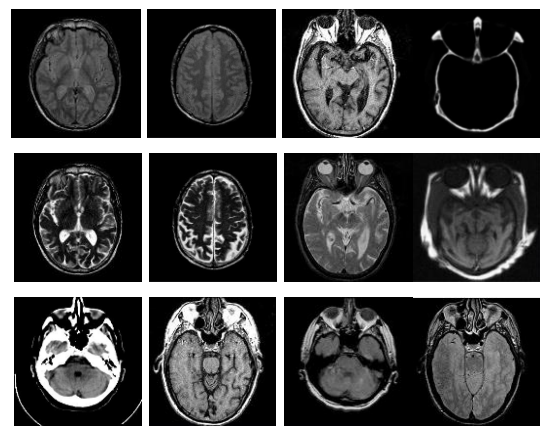
T1، T2 و PD را شامل می‌گردند، استفاده شده است. اندازه تمامی تصویرهای پایگاه داده‌های Brainweb برابر 217×181 و تمامی تصویرهای پایگاه داده‌های Harvard برابر 256×256 هستند.

تصویرهای چندکانونی: از ۱۰ جفت تصویر متداول به-کار رفته در مقالات، استفاده شده است. این تصویرها اندازه‌های متفاوتی دارند [۴۷].

تصویرهای چندحالتی مادون قرمز و مرئی: این تصویرها مربوط به پایگاه داده‌ها TNO [۴۸] هستند. تعداد ۹ جفت تصویر متداول از این پایگاه داده‌ها مورد استفاده قرار گرفته است. یک جفت تصویر متداول در مقالات، که شامل یک تصویر مرئی و یک تصویر موج میلیمتری (MMW) است نیز، در این دسته مقایسه‌ها استفاده شده است.



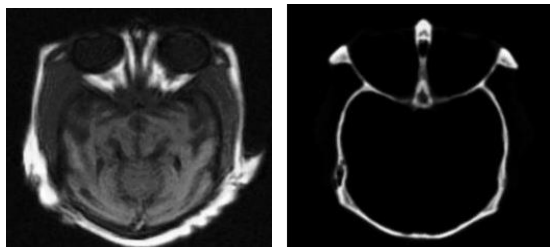
(الف) تصویرهای پایگاه داده Brainweb [۳۹]



(ب) تصویرهای پایگاه داده Harvard [۴۰]

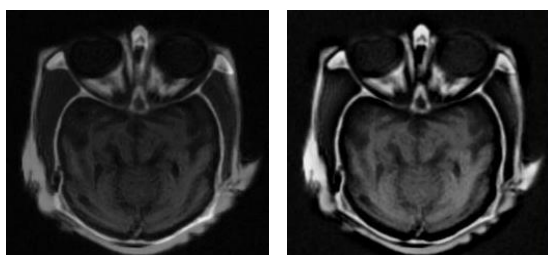


دوم قرار گرفته است. بنابراین بر اساس مجموعه معیارهای ارزیابی، روش پیشنهادی در مجموع بهترین خروجی را ارائه می‌دهد. ضعیف‌ترین عملکردهای کمی نیز در روش‌های CV، DTCWT و MGFF دیده می‌شوند.



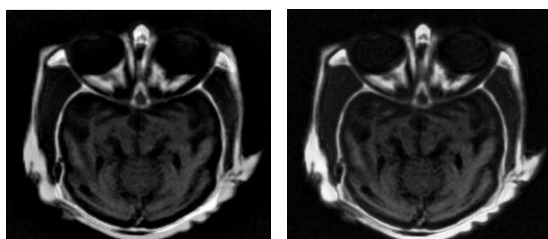
تصویر اولیه MRI

تصویر اولیه CT



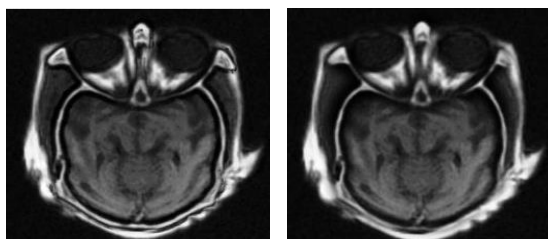
DTCWT (ب)

CVT (الف)



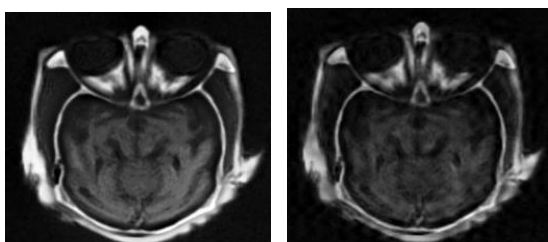
MGF (ت)

GFF (پ)



PSCBF (ج)

FFIF (ث)



NSCT-SPDF (ح)

DCHWT (چ)

شکل ۷- نمونه‌ای از تصویرهای ترکیب تصویرهای چندحالتی پزشکی حاصل از اعمال روش‌های مختلف

می‌آیند. در پیاده سازی NSCT، از فیلترهای هرمی "dmaxflat7" و فیلترهای جهتی "pyrexc" استفاده شده است.

۳-۴- نتایج شبیه سازی

مقایسه‌های کیفی به صورت بصری و مقایسه‌های کمی بر اساس معیارهای ارزیابی قید شده در بخش‌های پیشین صورت می‌گیرد. مقایسه‌ها بر اساس نوع تصویرهای ورودی، در سه دسته کلی ترکیب تصویرهای چندحالتی پزشکی، ترکیب تصویرهای چندکانونی و ترکیب تصویرهای مادون قرمز و مرئی انجام شده‌اند. برای محاسبه هر یک از معیارهای کمی، عمل ترکیب بر روی تمامی ۱۰ جفت تصویر، به صورت مجزا صورت گرفته و اندازه هر معیار، از میانگین معیارها برای ۱۰ جفت آزمایشی مذکور، حاصل می‌گردد.

۳-۴-۱- ترکیب تصویرهای چندحالتی پزشکی

نمونه‌ای از ترکیب دو تصویر CT و MRI در شکل (۷) نشان داده شده است. همان‌گونه که مشاهده می‌گردد روش‌های DTCWT، PSCBF و DCHWT به دلیل عدم حفظ کلیت شدت روشنایی تصویرهای اولیه و گسستگی‌های نامتعارف در برخی قسمت‌های تصویر، کیفیت بصری پایینی دارند. در خصوص CVT و MGF نیز در قسمت‌هایی که لبه‌ها و نواحی دارای شدت روشنایی بالاتر، در دو تصویر با یکدیگر ترکیب شده‌اند، تناسب خوبی بین شدت روشنایی‌ها در تصویر ترکیب شده دیده نمی‌شود. به‌طور کلی، تصویر خروجی حاصل از روش پیشنهادی، به خوبی قسمت‌های اثرگذار و حاوی اطلاعات پر اهمیت از دو تصویر را با یکدیگر ترکیب کرده و به‌طور هم‌زمان نشان می‌دهد. ضعف‌هایی مانند گسستگی نامتعارف و همچنین عدم تناسب شدت روشنایی‌ها در نواحی منطبق بر یکدیگر، به‌خوبی مرتفع شده و با توجه به حفظ لبه‌های تصویر و تفکیک‌پذیری مناسب، کیفیت بصری تصویر نهایی نسبت به دیگر روش‌ها بالاتر است.

جدول ۱ نشان دهنده مقایسه کمی ترکیب تصویرهای چندحالتی پزشکی است. ارزیابی‌ها عموماً منطبق بر مقایسه‌های کیفی صورت گرفته در شکل (۷) هستند. با توجه به کیفیت بصری مطلوب، روش پیشنهادی در سه معیار Q_G ، Q_{CV} و SF بهترین عملکرد را در مقایسه با روش‌های دیگر ارائه می‌دهد. در سه معیار دیگر نیز، روش پیشنهادی در بین تمامی روش‌های ارزیابی شده در رتبه

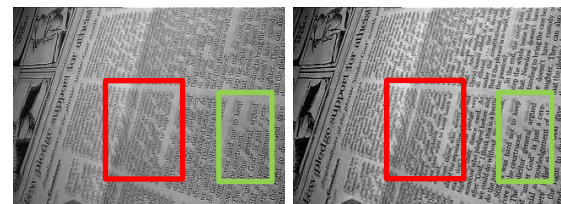
۳-۴-۲- ترکیب تصویرهای چندکانونی

شکل (۸)، نمونه‌ای از ترکیب دو تصویر با دو کانون مختلف را نشان می‌دهد.



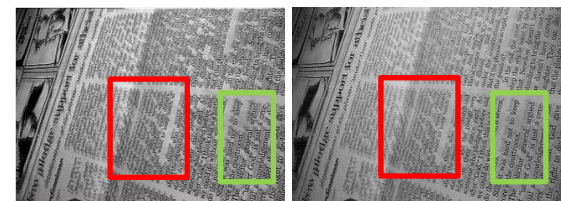
تصویر اولیه B

تصویر اولیه A



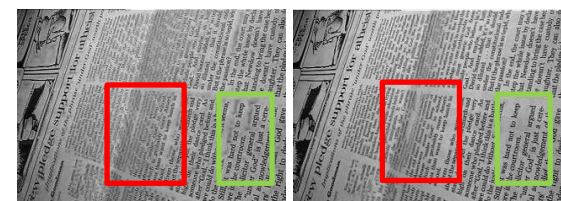
DTCWT (ب)

CVT (الف)



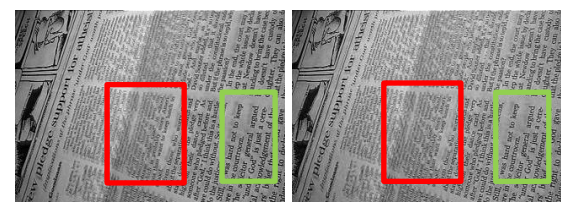
MGF (ت)

GFF (پ)



PSCBF (ج)

FFIF (ث)



NSCT-SPDF (ح)

DCHWT (چ)

شکل ۸- نمونه‌ای از تصویرهای ترکیب خروجی حاصل از اعمال روش‌های مختلف برای تصویرهای چندکانونی

همان‌گونه که مشاهده می‌گردد روش‌های DTCWT، MGF، GFF و تا حدودی CVT، منجر به خروجی تار و درهم ریخته‌ای می‌شوند که کیفیت بصری پایینی دارند.

پدیده نامطلوب حلقه شدن^۱، خصوصاً در قسمت‌هایی که متن روزنامه وجود دارد، در اکثر روش‌ها دیده می‌شود. در مجموع، کیفیت بصری تصویر ترکیب شده FFIF از بقیه روش‌ها بهتر است و روش پیشنهادی نیز از جمله روش‌های با خروجی بصری مناسب در این نمونه تصویر آزمایشی است.

مقایسه‌های کمی برای ترکیب تصویرهای چندکانونی در جدول ۲ نشان داده شده است. روش پیشنهادی، در چهار معیار Q_S ، Q_{SF} ، Q_G و Q_P بهترین عملکرد را دارد؛ در معیار Q_P در رده دوم و در Q_{CV} در سومین رده قرار می‌گیرد. بنابراین روش پیشنهادی، بر اساس مجموعه معیارهای ارزیابی کمی، به‌طور کلی بهترین عملکرد را ارائه می‌دهد. ضعیف‌ترین عملکردهای کمی نیز همانگونه که در نمونه شکل (۸) مشاهده می‌شود متعلق به روش‌های DTCWT، CVT، GFF و MGF است.

۳-۴-۳- ترکیب تصویرهای مرئی و مادون قرمز

در شکل (۹)، نمونه‌ای از ترکیب دو تصویر مرئی و مادون قرمز ارائه شده است. همان‌گونه که مشاهده می‌گردد در روش CVT ماهیت تصویر خصوصاً در قسمت پس‌زمینه درهم ریخته و نسبت به تصویرهای اولیه ورودی تغییراتی دارد. در روش‌های FFIF، PSCBF و DCHWT، پس‌زمینه و ساختمان پایگاه وسط تصویر، قابل تفکیک نیستند. DTCWT نیز از تفکیک‌پذیری مناسبی برخوردار نیست. در مجموع به لحاظ بصری، روش پیشنهادی علاوه بر تفکیک‌پذیری مطلوب، از تفکیک‌پذیری خوبی بین محیط پس‌زمینه و تصویر پایگاه برخوردار است و به نوعی بهترین کیفیت بصری را ارائه می‌دهد.

در جدول ۳، مقایسه روش پیشنهادی با دیگر روش‌های مدنظر، ارائه شده است. در این جدول نیز، روش پیشنهادی در سه معیار Q_P ، Q_S و Q_G بهترین عملکرد را ارائه می‌دهد و در دو معیار Q_{SF} و Q_{CV} نیز در رده دوم قرار دارد. با توجه به مجموعه معیارهای مورد ارزیابی، روش پیشنهادی در مجموع، عملکرد بهتری نسبت به تمامی روش‌ها ارائه می‌کند. روش‌های CVT و DTCWT نیز ضعیف‌ترین عملکرد را نشان می‌دهند.

۴-۴- پیچیدگی محاسباتی و زمان اجرا

اگر تعداد تصویرهای اولیه که قصد ترکیب هم‌زمان آنها را

¹ Ringing

داریم برابر K و تعداد کل پیکسل‌های هر تصویر اولیه M باشد، با توجه به اینکه NSCT شامل P سطح و D جهت است، پیچیدگی روش پیشنهادی برابر $O(MKP(2^D))$ است. جدول ۴ مقایسه‌ای از میانگین زمان اجرای روش‌های مختلف را به تفکیک دسته‌بندی تصویرهای مختلف نشان می‌دهد. شبیه‌سازی‌ها در کامپیوتری با پردازشگر Intel core i5@2.5GHz و حافظه RAM برابر ۴ گیگابایت صورت گرفته است. برای مقایسه عادلانه زمان اجرا، از کد اصلی مقالات که به صورت عمومی در اختیار است استفاده شده است.

به طور میانگین، اندازه تصویرهای چندحالتی پزشکی مورد استفاده در بررسی‌ها، از دو دسته دیگر کمتر است. تصویرهای چندکانونی از این لحاظ در رده دوم قرار می‌گیرند و میانگین اندازه مجموعه ۱۰ تصویر مادون قرمز و مرئی استفاده شده، بیشترین است و در واقع عموماً تصویرهای بزرگتری را شامل می‌شوند. به همین دلیل نیز مدت زمان اجرا به ترتیب برای تصویرهای چندحالتی پزشکی کمترین و برای تصویرهای مرئی و مادون قرمز بیشترین است. با توجه به نتایج، FFIF بهترین عملکرد را داشته و سریع‌ترین روش است. روش پیشنهادی نیز به لحاظ زمان اجرا در رتبه چهارم قرار داشته و به نوعی عملکرد متوسطی ارائه می‌دهد. همان‌گونه که از رابطه پیچیدگی مشخص است، استفاده از NSCT هم در تعداد سطوح (P) و هم تعداد جهت‌ها (D) اثرگذار است. با افزایش تعداد سطوح و جهت‌ها، کیفیت ترکیب تصاویر افزایش می‌یابد ولی در ازای آن، پیچیدگی بیشتر شده و زمان بیشتری مورد نیاز است که خصوصاً در کاربردهای زمان واقعی بسیار اهمیت دارد.

با توجه به مصالحه‌ای که بین زمان و کیفیت وجود دارد، برای رسیدن به کیفیت مطلوب به عنوان هدف اصلی (در عین در نظر داشتن زمان اجرا به عنوان هدف ثانویه)، در تبدیل NSCT از $P = 4$ و $D = 1$ استفاده شده است.



تصویر اولیه مرئی

تصویر اولیه مادون قرمز



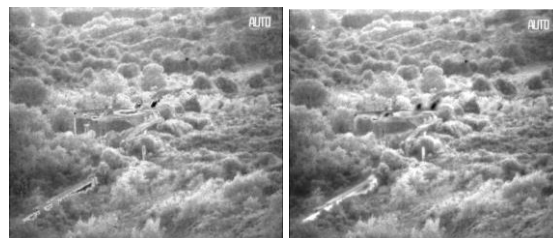
DTCWT (ب)

CVT (الف)



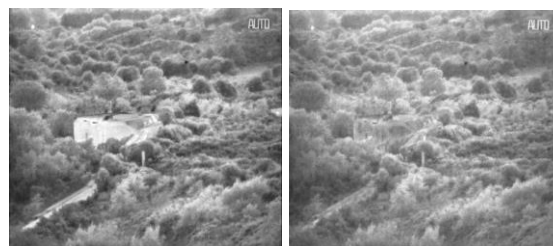
MGF (ت)

GFF (پ)



PSCBF (ج)

FFIF (ث)



NSCT-SPDF (ح)

DCHWT (چ)

شکل ۹- نمونه‌ای از تصویرهای ترکیب خروجی حاصل از اعمال روش‌های مختلف برای تصویرهای مادون قرمز و مرئی

جدول ۱- مقایسه کمی ترکیب روش پیشنهادی و دیگر روش‌های متداول، برای تصویرهای چندحالتی پزشکی (در هر سطر، بهترین معیار توپیر شده است).

معیار	CVT	DTCWT	GFF	MGF	FFIF	PSCBF	DCHWT	NSCT-SPDF
Q_G	۰/۴۸۵۶	۰/۵۲۱۷	۰/۶۱۶۶	۰/۵۸۹۱	۰/۶۷۱۶	۰/۶۹۶۶	۰/۶۱۳۹	۰/۶۹۷۸
Q_{SF}	-۰/۳۱۰۶	-۰/۳۰۲۸	-۰/۲۲۳	-۰/۱۰۴۱	-۰/۲۱۳۹	-۰/۱۱۷۴	-۰/۲۱۸۹	-۰/۱۱۶
Q_P	۰/۴۱۲۹	۰/۴۴۶۶	۰/۴۶۳۰	۰/۳۶۳۵	۰/۶۲۰۴	۰/۴۶۹۴	۰/۴۶۴۰	۰/۵۳۷۲
Q_S	۰/۶۹۰۳	۰/۷۰۴۲	۰/۷۷۹۷	۰/۷۰۱۷	۰/۷۲۸۸	۰/۸۱۵۷	۰/۷۸۲۲	۰/۷۹۳۲
Q_{CV}	۱۱۷۹/۲۹	۱۱۱۷/۷۱۵	۷۴۲/۷۸	۸۸۷/۴۱	۱۴۲۸/۴۳	۷۷۲/۰۸	۷۴۷/۶۸	۷۱۱/۶۳
SF	۲۸/۰۷۴	۲۸/۳۰۹	۳۰/۶۸۷	۳۳/۶۱۸	۳۰/۶۰۰	۳۴/۵۷۴	۳۱/۱۴۱	۳۴/۷۵۵

جدول ۲- مقایسه کمی ترکیب روش پیشنهادی و دیگر روش‌های متداول، برای تصویرهای چندکانونی (در هر سطر، بهترین معیار توپیر شده است).

معیار	CVT	DTCWT	GFF	MGF	FFIF	PSCBF	DCHWT	NSCT-SPDF
Q_G	۰/۶۳۹۷	۰/۶۷۲۵	۰/۶۷۸۷	۰/۶۷۸۹	۰/۷۴۶۷	۰/۷۴۶۴	۰/۷۱۹۰	۰/۷۴۹۶
Q_{SF}	-۰/۲۴۵۹	-۰/۲۳۷۵	-۰/۲۹۴۱	-۰/۱۶۲۵	-۰/۰۹۵۱	-۰/۰۸۹۹	-۰/۱۲۳۵	-۰/۰۵۹۰
Q_P	۰/۶۵۶۵	۰/۶۹۶۳	۰/۷۰۸۰	۰/۶۳۱۸	۰/۷۹۷۰	۰/۷۴۴۱	۰/۷۱۳۳	۰/۷۶۶۸
Q_S	۰/۸۷۱۲	۰/۸۸۵۱	۰/۸۷۸۸	۰/۸۹۱۹	۰/۸۸۶۰	۰/۹۱۸۹	۰/۹۱۷۳	۰/۹۲۰۲
Q_{CV}	۱۰۳/۲۱	۹۲/۸۷۹	۶۹/۳۲۰	۸۶/۱۶۸	۶۷/۵۲۷	۳۹/۸۸۰	۴۸/۸۴۹	۵۴/۶۹۳
SF	۱۸/۴۲۳	۱۸/۴۵۵	۱۶/۱۴۷	۱۹/۰۶۹	۲۱/۶۷۷	۲۰/۹۲۷	۲۰/۴۷۷	۲۱/۷۷۰

جدول ۳- مقایسه کمی ترکیب روش پیشنهادی و دیگر روش‌های متداول، برای تصویرهای مرئی و مادون قرمز (در هر سطر، بهترین معیار توپیر شده است).

معیار	CVT	DTCWT	GFF	MGF	FFIF	PSCBF	DCHWT	NSCT-SPDF
Q_G	۰/۴۴۰۷	۰/۴۷۸۶	۰/۵۹۶۱	۰/۵۸۱۹	۰/۶۳۲۹	۰/۵۵۹۵	۰/۵۵۹۱	۰/۶۶۸۳
Q_{SF}	-۰/۲۶۸۴	-۰/۲۸۸۳	-۰/۲۲۶۸	-۰/۰۴۵۲	-۰/۱۰۸۶	-۰/۱۱۳۵	-۰/۱۸۴۵	-۰/۰۸۰۱
Q_P	۰/۳۰۳۰	۰/۳۲۶۰	۰/۳۹۸۸	۰/۳۱۴۰	۰/۴۰۲۶	۰/۲۱۷۶	۰/۲۷۸۱	۰/۴۱۳۳
Q_S	۰/۷۵۵۰	۰/۷۶۵۹	۰/۸۳۲۲	۰/۸۲۴۷	۰/۸۰۸۰	۰/۸۳۴۹	۰/۸۵۴۹	۰/۸۶۸۹
Q_{CV}	۵۶۱/۴۹	۵۶۱/۷۵	۳۳۳/۲۱	۵۱۱/۸۷	۱۰۵۰/۴۷	۸۴۱/۱۱	۴۴۸/۴۵	۴۳۹/۴۸
SF	۹/۲۷۰۳	۹/۰۸۸۱	۹/۶۸۴۱	۱۱/۹۹۹	۱۱/۹۶۳	۱۳/۴۱۴	۱۰/۲۷۳	۱۱/۵۳۴

جدول ۴- مقایسه زمان اجرای روش پیشنهادی با دیگر روش‌های متداول. تمامی مقادیر بر حسب ثانیه هستند. (در هر سطر، بهترین معیار توپر شده است.)

NSCT-SPDF	DCHWT	PSCBF	FFIF	MGF	GFF	DTCWT	CVT	
۰/۴۳۹۳	۱/۸۵	۲۲/۸۶۲	۰/۰۲۹۹	۰/۰۷۳۱	۰/۰۷۰۳	۰/۱۳۶۳	۰/۴۵۰۵	تصویرهای چندحالتی پزشکی
۱/۸۸	۵/۴۱۸	۶۶/۸۷	۰/۱۰۳۶	۰/۴۹۴	۰/۴۳۴۱	۰/۲۰۱۳	۱/۱۱	تصویرهای چندکانونی
۲/۹۵۵	۷/۳۶۵	۱۰۱/۹۰	۰/۱۹۲۴	۰/۷۸۸۲	۰/۷۱۸۱	۰/۴۸۲۶	۳/۲۰۸	تصویرهای مرئی و مادون قرمز
۱/۷۵۸۱	۴/۸۷۸	۶۳/۸۷۷	۰/۱۰۸۶	۰/۴۵۱۷	۰/۴۰۷۵	۰/۲۷۳	۱/۵۸۹	میانگین هر سه دسته

جدول ۵- مقایسه کمی ترکیب روش پیشنهادی و دیگر روش‌های متداول. مقادیر، حاصل از میانگین جدول‌های ۱ تا ۳ هستند. (در هر سطر، بهترین معیار توپر شده است.)

NSCT-SPDF	DCHWT	PSCBF	FFIF	MGF	GFF	DTCWT	CVT	معیار
۰/۷۰۴۷	۰/۶۳۰۳	۰/۶۶۷۰	۰/۶۸۳۰	۰/۶۱۶۰	۰/۶۳۰۰	۰/۵۵۷۰	۰/۵۲۱۳	Q_G
-۰/۰۸۵۰	-۰/۱۷۵۲	-۰/۱۰۶۳	-۰/۱۳۸۷	-۰/۱۰۳۷	-۰/۲۴۷۷	-۰/۲۷۵۷	-۰/۲۷۴۳	Q_{SF}
۰/۵۷۲۰	۰/۴۸۵۰	۰/۴۷۶۷	۰/۶۰۶۳	۰/۴۳۶۰	۰/۵۲۳۰	۰/۴۸۹۳	۰/۴۵۷۰	Q_P
۰/۸۶۰۳	۰/۸۵۱۰	۰/۸۵۵۷	۰/۸۰۷۳	۰/۸۰۵۳	۰/۸۲۹۷	۰/۷۸۴۷	۰/۷۷۲۰	Q_S
۴۰۱/۹۳	۴۱۴/۹۹	۵۵۱/۰۲	۸۴۸/۸۰	۴۹۵/۱۴	۳۸۱/۷۷	۵۹۰/۷۷	۶۱۴/۶۶	Q_{CV}
۲۲/۶۸۳	۲۰/۶۲۶	۲۲/۵۶۶	۲۱/۴۱۰	۲۱/۵۵۳	۱۸/۸۳۳	۱۸/۶۱۰	۱۸/۵۸۶	SF

است، بیانگر عملکرد کلی روش پیشنهادی در مقایسه با روش‌های دیگر است. از مجموع ۶ معیار مورد ارزیابی، در چهار معیار Q_G ، Q_{SF} ، Q_S و Q_P ، روش پیشنهادی بهترین عملکرد را نسبت به ۷ روش متداول دیگر نشان داده است و در دو معیار Q_{CV} و Q_P نیز در رتبه دوم قرار می‌گیرد. این نتایج موید برتری روش پیشنهادی نسبت به روش‌های دیگر مورد ارزیابی است. بنابراین بهره‌گیری از تبدیل NSCT که علاوه بر تقریب، جزئیات را نیز در جهت‌ها و مقیاس‌های مختلف مورد توجه قرار می‌دهد و بهره‌گیری از بازنمایی SPD و اعمال هرمی آن با استفاده از ضریب وزنی

۵-۴- جمع‌بندی ارزیابی‌ها

آزمایش‌های مختلف و بررسی‌های انجام شده در بخش‌های پیشین (شکل‌های (۹-۷) و جدول‌های (۴-۱)) نشان می‌دهد که در ترکیب هر سه دسته تصویرهای چندحالتی پزشکی، چندکانونی و مرئی و مادون قرمز، روش پیشنهادی در مجموع هم از لحاظ بصری و هم معیارهای ارزیابی کمی، عملکرد بهتری نسبت به تمامی روش‌ها ارائه می‌کند. در هر سه دسته تصاویر، حداقل در ۳ معیار از مجموع ۶ معیار بررسی شده، روش پیشنهادی بهترین عملکرد را نشان می‌دهد. جدول ۵ که حاصل از میانگین جدول‌های ۱ تا ۳

انجام شده و تصویر ترکیب شده نهایی (حاصل از تصویرهای اولیه) به دست می آید.

در روند تکراری ترکیب تعداد K زیرباند متناظر، برای هر یک از تصویرهای این زیرباندها، ابتدا ساختار هرمی مجزایی ایجاد می شود که شامل لایه های تقریب و جزئیات است. ایجاد لایه های مذکور، بر اساس نمونه برداری کاهش با ضریب ۲ و SPD ای که در آن از ضریب وزنی پیشنهادی جدید استفاده شده است، صورت می گیرد. با استفاده از این ضریب وزنی، جزئیات بیشتری از تصویر، مدنظر قرار گرفته و در تصویر ترکیب شده نهایی ظاهر می شوند. عمل ترکیب، در جهت عکس ساختار هرمی و با استفاده از نمونه برداری افزایشی و ترکیب بخش های مختلف SPD صورت می گیرد. تصویر حاصل، در واقع تصویر ترکیب شده حاصل از K زیرباند متناظر است. با تکرار این روند، تصویر ترکیب شده ی تمامی زیرباندهای متناظر، به دست آمده و با اعمال NSCT معکوس، تصویر ترکیب شده نهایی حاصل می شود.

ارزیابی های کمی و بصری، در مقایسه با ۷ روش متداول و بر اساس ۶ معیار کمی مختلف صورت گرفته است. این مقایسه ها برای سه دسته تصاویر مجزا شامل ترکیب تصویرهای چندحالتی پزشکی، ترکیب تصویرهای چندکانونی و ترکیب تصویرهای مرئی و مادون قرمز انجام شده است. در هر یک از سه دسته مقایسه مجزای صورت گرفته، علاوه بر اینکه روش پیشنهادی بالاترین کیفیت بصری را داشته است، به لحاظ معیارهای کمی نیز حداقل در ۳ معیار بهترین بوده و در بقیه معیارها نیز از جمله بهترین ها بوده است. در مقایسه ای که از میانگین مقادیر هر سه دسته حاصل شده نیز، از بین ۶ معیار ذکر شده، روش پیشنهادی در ۴ معیار بهترین بوده است. روش پیشنهادی از لحاظ زمان و پیچیدگی محاسبات نیز، دارای عملکرد متوسطی بوده و به طور متوسط $1/75$ ثانیه برای ترکیب دو تصویر زمان صرف می کند. ترکیب تصویرهای سایر حوزه های تصویربرداری و همچنین ترکیب مستقیم تصویرهای سه بعدی حوزه پزشکی را می توان در پژوهش های آتی، مدنظر قرار داد.

پیشنهادی که موجب پرداختن و تاکید بیشتر بر جزئیات می شود، مطابق انتظار، منجر به بهبود عملکرد نسبت به دیگر روش های متداول شده است.

علیرغم اینکه هدف اصلی در روش پیشنهادی، دستیابی به کیفیت بالاتر در تصویر ترکیب شده است، این روش از لحاظ مدت زمان اجرا نیز عملکرد متوسطی دارد. بر اساس مدت زمان اجرا و پیچیدگی محاسباتی، امکان استفاده از روش پیشنهادی در عموم کاربردهای زمان واقعی وجود دارد. به عنوان بررسی های آتی، بهره گیری از پیش پردازش های اولیه از جمله حذف نویز، می تواند به بهبود کیفیت تصویر خروجی و مقاوم سازی آن کمک کند [۴۹].

۵- نتیجه گیری

در این مقاله، روشی برای ترکیب تصویرها، مبتنی بر تفکیک ساختاری تصویرها و با بهره گیری از NSCT و ساختار هرمی پیشنهاد شده است. تفکیک ساختاری SPD، اخیراً برای استفاده در ترکیب تصویرهای چندنوری معرفی شده و تصویرها را به سه بخش مختلف شامل قدرت سیگنال، ساختار سیگنال و میانگین سیگنال، تفکیک می کند. در روش پیشنهادی این مقاله، با افزودن NSCT و معرفی ضریب وزنی جدید، از SPD برای ترکیب انواع دیگری از تصویرها که شامل تصویرهای چندحالتی پزشکی، تصویرهای چندکانونی و تصویرهای مادون قرمز و مرئی هستند، استفاده شده است. استفاده از تبدیل NSCT، به دلیل پوشش دادن ساختارهای پیچیده و با مرتبه بالاتر (لبه ها، جهت ها و جزئیات)، باعث بهبود جزئیات و کیفیت تصویر خروجی می گردد.

در روش پیشنهادی، تعداد K تصویر اولیه، با استفاده از NSCT (در مقیاس ها و جهت های مختلف)، تبدیل به تصویرهای زیرباندهای فرکانس پایین و فرکانس بالا می شوند. در ادامه، هر مجموعه K تایی از زیرباندهایی که متناظر هستند (زیرباندهای حاصل از مقیاس و جهت یکسان)، به صورت مجزا و با استفاده از یک روند تکراری، با یکدیگر ترکیب می شوند. در انتها، پس از به دست آوردن تمامی زیرباندهای ترکیب شده، عمل NSCT معکوس

مراجع

- [1] Kaur, H., Koundal, D., & Kadyan, V., "Image fusion techniques: a survey." Archives of computational methods in Engineering 28, no. 7, 2021, pp. 4425-4447.
- [2] Azam, M. A., Khan, K. B., Salahuddin, S., Rehman, E., Khan, S. A., Khan, M. A. & Gandomi, A. H., "A review on multimodal medical image fusion: Compendious analysis of medical modalities, multimodal databases, fusion techniques and quality metrics." Computers in Biology and Medicine 144, 2022, pp. 105253.
- [3] Song L, Lin Y, Feng W, Zhao M , "A novel automatic weighted image fusion algorithm." In 2009 International Workshop on Intelligent Systems and Applications, IEEE, 2009, pp. 1-4.
- [4] Singh, H., Kumar, V., & Bhooshan, S., "Weighted least squares based detail enhanced exposure fusion." International Scholarly Research Notices, 2014.
- [5] Bavirisetti, D. P., Xiao, G., & Liu, G., "Multi-sensor image fusion based on fourth order partial differential equations." In 2017 20th International conference on information fusion (Fusion), IEEE, 2017, pp. 1-9.
- [6] Zhang, X., Lin, H., Kang, X., & Li, S., "Multi-modal image fusion with KNN matting." In Chinese Conference on Pattern Recognition, Springer, Berlin, Heidelberg, 2014. pp. 89-96.
- [7] Shreyamsha Kumar, B. K., "Image fusion based on pixel significance using cross bilateral filter." Signal, image and video processing 9, no. 5, 2015, pp. 1193-1204.
- [8] Zhan, K., Xie, Y., Wang, H., & Min, Y., "Fast filtering image fusion." Journal of Electronic Imaging 26, no. 6, 2017, pp. 063004.
- [9] Ma, J., & Zhou, Y., "Infrared and visible image fusion via gradientlet filter." Computer Vision and Image Understanding 197, 2020, pp. 103016.
- [10] Zhang, Y., Wang, M., Xia, X., Sun, D., Zhou, X., Wang, Y., & Huang, G., "Medical image fusion based on quasi-cross bilateral filtering." Biomedical Signal Processing and Control, 80, 2023, pp. 104259.
- [11] Zhang, X., Dai, X., Zhang, X., & Jin, G., "Joint Principal Component Analysis and Total Variation for Infrared and Visible Image Fusion." Infrared Physics & Technology, 2022, pp. 104523.
- [12] Yang, Y., Cao, S., Wan, W., & Huang, S., "Multi-modal medical image super-resolution fusion based on detail enhancement and weighted local energy deviation." Biomedical Signal Processing and Control, 80, 2023, pp. 104387.
- [13] Li, S., Kang, X., & Hu, J., "Image fusion with guided filtering." IEEE Transactions on Image processing 22, no. 7, 2013, pp. 2864-2875.
- [14] Naidu, V. P. S., "Image fusion technique using multi-resolution singular value decomposition." Defence Science Journal 61, no. 5, 2011, pp. 479.
- [15] Zhan, K., Li, Q., Teng, J., Wang, M., & Shi, J., "Multifocus image fusion using phase congruency." Journal of Electronic Imaging 24, no. 3, 2015, pp. 033014.
- [16] Hayat, N., & Imran, M., "Ghost-free multi exposure image fusion technique using dense SIFT descriptor and guided filter." Journal of Visual Communication and Image Representation 62, 2019, pp. 295-308.
- [17] Zhang, Q., & Guo, B. L., "Multifocus image fusion using the nonsubsampling contourlet transform." Signal processing 89, no. 7, 2009, pp. 1334-1346.
- [18] Li, X., Wan, W., Zhou, F., Cheng, X., Jie, Y., & Tan, H., "Medical image fusion based on sparse representation and neighbor energy activity." Biomedical Signal Processing and Control, 80, 2023, pp. 104353.
- [19] Diwakar, M., Singh, P., & Shankar, A., "Multi-modal medical image fusion framework using co-occurrence filter and local extrema in NSST domain." Biomedical Signal Processing and Control, 68, 2021, pp. 102788.
- [20] Nencini, F., Garzelli, A., Baronti, S., & Alparone, L., "Remote sensing image fusion using the curvelet transform." Information fusion 8, no. 2, 2007, pp. 143-156.
- [21] Lewis, J. J., O'Callaghan, R. J., Nikolov, S. G., Bull, D. R., & Canagarajah, N., "Pixel-and region-based image fusion with complex wavelets." Information fusion 8, no. 2, 2007, pp. 119-130.
- [22] Shreyamsha Kumar, B. K., "Multifocus and multispectral image fusion based on pixel significance using discrete cosine harmonic wavelet transform." Signal, Image and Video Processing 7, no. 6, 2013, pp. 1125-1143.

- [23] Bavirisetti, D. P., Xiao, G., Zhao, J., Dhuli, R., & Liu, G., "Multi-scale guided image and video fusion: A fast and efficient approach." *Circuits, Systems, and Signal Processing* 38, no. 12, 2019, pp. 5576-5605.
- [24] Li, H., Ma, K., Yong, H., & Zhang, L., "Fast multi-scale structural patch decomposition for multi-exposure image fusion." *IEEE Transactions on Image Processing* 29, 2020, pp. 5805-5816.
- [25] Ma, K., Li, H., Yong, H., Wang, Z., Meng, D., & Zhang, L., "Robust multi-exposure image fusion: a structural patch decomposition approach." *IEEE Transactions on Image Processing*, 26, no. 5, 2017, pp. 2519-2532.
- [26] Zhang, H., Xu, H., Tian, X., Jiang, J., & Ma, J., "Image fusion meets deep learning: A survey and perspective." *Information Fusion* 76, 2021, pp. 323-336.
- [۲۷] مرضیه زارع نظری، محسن سرداری زارچی، سیما عمادی و هادی پورمحمدی، "چارچوب برای استخراج آناتومی و طبقه بندی تصاویر پشه با رویکرد یادگیری عمیق"، نشریه مدل سازی در مهندسی، دوره ۲۰، شماره ۷۰، بهار ۱۴۰۱.
- [۲۸] مسلم سردشتی بیرجندی، حسین رحمانی و سعید فراهت، "کاربرد شبکه های عصبی عمیق در طبقه بندی تصاویر آسیب های شبکه فاضلاب و مشخص کردن مسیرهای بحرانی آنها"، نشریه مدل سازی در مهندسی، دوره ۲۰، شماره ۷۰، بهار ۱۴۰۱.
- [29] Amin-Naji, M., Aghagolzadeh, A., & Ezoji, M., "Ensemble of CNN for multi-focus image fusion." *Information fusion* 51, 2019, pp. 201-214.
- [30] Gao, Y., Ma, S., Liu, J., & Xiu, X., "Fusion-UDCGAN: Multifocus Image Fusion via a U-Type Densely Connected Generation Adversarial Network." *IEEE Transactions on Instrumentation and Measurement*, 71, 2022, pp. 1-13.
- [31] Xia, K. J., Yin, H. S., & Wang, J. Q., "A novel improved deep convolutional neural network model for medical image fusion." *Cluster Computing* 22, no. 1, 2019, pp. 1515-1527.
- [32] Zhang, B., Jiang, C., Hu, Y., & Chen, Z., "Medical image fusion based a densely connected convolutional networks." In *2021 IEEE 5th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*, IEEE, 2021, vol. 5, pp. 2164-2170
- [33] Zhang, H., Yuan, J., Tian, X., & Ma, J., "GAN-FM: Infrared and visible image fusion using GAN with full-scale skip connection and dual Markovian discriminators." *IEEE Transactions on Computational Imaging* 7, 2021, pp. 1134-1147.
- [34] Upla, K. P., Joshi, M. V., & Gajjar, P. P., "An edge preserving multiresolution fusion: Use of contourlet transform and MRF prior." *IEEE Transactions on Geoscience and Remote Sensing* 53, no. 6, 2014, pp. 3210-3220.
- [35] Zhang, Q., & Guo, B. L., "Multifocus image fusion using the nonsubsampling contourlet transform." *Signal processing* 89, no. 7, 2009, pp. 1334-1346.
- [36] Da Cunha, A. L., Zhou, J., & Do, M. N., "The nonsubsampling contourlet transform: theory, design, and applications." *IEEE transactions on image processing* 15, no. 10, 2006, pp. 3089-3101.
- [37] He, K., Sun, J., & Tang, X., "Guided image filtering." *IEEE transactions on pattern analysis and machine intelligence* 35, no. 6, 2012, pp. 1397-1409.
- [38] Yin, M., Liu, X., Liu, Y., & Chen, X., "Medical image fusion with parameter-adaptive pulse coupled neural network in nonsubsampling shearlet transform domain." *IEEE Transactions on Instrumentation and Measurement* 68, no. 1, 2018, pp. 49-64.
- [39] Xydeas, C. S., & Petrovic, V., "Objective image fusion performance measure." *Electronics letters* 36, no. 4, 2000, pp. 308-309.
- [40] Zheng, Y., Essock, E. A., Hansen, B. C., & Haun, A. M., "A new metric based on extended spatial frequency and its application to DWT based fusion algorithms." *Information Fusion* 8, no. 2, 2007, pp. 177-192.
- [41] Zhao, J., Laganieri, R., & Liu, Z., "Performance assessment of combinative pixel-level image fusion based on an absolute feature measurement." *Int. J. Innov. Comput. Inf. Control* 3, no. 6, 2007, pp. 1433-1447.
- [42] Piella, G., & Heijmans, H., "A new quality metric for image fusion." In *Proceedings 2003 international conference on image processing IEEE*, (Cat. No. 03CH37429), 2003, vol. 3, pp. III-173.
- [43] Chen, H., & Varshney, P. K., "A human perception inspired quality metric for image fusion based on regional information." *Information fusion* 8, no. 2, 2007, pp. 193-207.
- [44] Li, S., Kwok, J. T., & Wang, Y., "Combination of images with diverse focuses using the spatial frequency." *Information fusion* 2, no. 3, 2001, pp. 169-176.

[45] Neurological, M. B. I. C. M. (2020). Institute, McGill University, Montreal, QC, Canada. BrainWeb.. <https://brainweb.bic.mni.mcgill.ca/brainweb/>. Accessed April. 1, 2022.

[46] Johnson, K. A., & Becker, J. A. (1999). The Whole Brain Atlas, Harvard Medical School. <https://www.med.harvard.edu/aanlib>. April. 1, 2022.

[47] Liu, Y., Wang, L., Cheng, J., Li, C., & Chen, X., "Multi-focus image fusion: A survey of the state of the art." *Information Fusion* 64, 2020, pp. 71-91.

[48] Alexander Toet et al. Tno image fusion dataset. <https://doi.org/10.6084/m9.figshare.1008029.v1> , April. 1, 2022.

[۴۹] اصغر زارع و علی محمدزاده، "حذف نویز ضربه‌ای از تصاویر دیجیتالی مبتنی بر تخمین توزیع مکانی نویزها"، نشریه مدل‌سازی در مهندسی، دوره ۱۲، شماره ۳۹، زمستان ۱۳۹۳، صفحه ۱۳-۲۹.