Complex-scalar and complex-vector approaches for express target-oriented image fusion

[1] Khaustov D. Ye., [1] Kyrychuk O. A., [1] Stakh T. M., [1] Khaustov Ya. Ye., [1] Burashnikov O. O., [1] Ryzhov Ye., [2] Vlokh R. and [1] Nastishin Yu. A.

[1] Hetman Petro Sahaidachnyi National Army Academy, 32 Heroes of Maidan Street, 79012 Lviv, Ukraine
[2] O. G. Vlokh Institute of Physical Optics, 23 Dragomanov Street, 79005 Lviv, Ukraine

Abstract. Express techniques of target-oriented image fusion (IF) aim at target-data acquisition such as visual search, detection, recognition and identification of a target. They form a distinct class of the IF methods different from those designed for art photography. The most evident argument for their distinction is that the requirements concerned with the complexity of image-processing procedures and image-quality assessment are essentially different for the two classes. Namely, these requirements are of primary importance for one of the above classes and of no significance (or even inapplicable) for the other class. The express target-oriented IF methods need their specific quantitative image-quality indices, which are selectively associated with target visibility and conspicuity. After discussing the specific features of assessment of the quality of images fused using the express target-oriented IF methods, we suggest a method for determining the local contrast of a target and relate this index to the visibility of this target. We focus on applying the complex-scalar IF and complex-vector IF (CVIF) approaches, which have been developed recently, to the problems of express target-oriented IF. It is demonstrated that the ellipticity algorithm of the CVIF provides a considerable enhancement of the target visibility in comparison with the other express target-oriented IF methods.

1. Introduction

Traditional definitions consider an image fusion (IF) as a combination of partial images of the same scene, which carry different information, into one fused image for enhancing its quality. For the sake of generality of this definition, the term 'image quality' is used regardless of the aim of IF. However, it is clear that what is understood under the image quality depends essentially on the IF aim. Indeed, what is characterized as a high-quality image in terms of amateur or professional art photography might be of a poor quality if the IF should facilitate acquisition of some target data (i.e., search, detection, recognition and identification of a target), and vice versa. In its vast majority, the art IF deals with the images obtained by cameras that work in the visible (Vis) light while, in most cases, the target-oriented data acquisition uses the images obtained in different spectral ranges of electromagnetic waves.

The partial images obtained using the same or similar cameras (usually conventional digital cameras) that work in the same spectral ranges (conventionally, in the Vis light) under different exposition conditions (time, viewing angle, lighting, zooming, etc.) are called *mono-modal* partial

images. The partial images obtained in different spectral ranges of electromagnetic waves are called *multi-modal* partial images. While the fusion of mono-modal partial images in the Vis light is interesting mostly for the art photography, the multi-modal partial images are fused for the activities embraced by the term '*visual search*' [1–3]. Its task is the target-data acquisition in the presence of irrelevant distracting (non-target) elements. Fusion of the multimodal partial images is needed because the human visual perception is poor at low lighting, leaving us blind in the absence of visual (Vis) light (i.e., the electromagnetic waves located in a relatively narrow spectral range, approximately between the wavelengths 350 and 750 nm). Special digital-matrix detectors are available nowadays to solve the problem of human blindness outside the Vis spectral range. The fusion of the multi-modal partial images collected in different ranges of the electromagnetic spectrum is aimed to improve the *visibility* and *conspicuity* of targets.

There are several dozens of IF techniques known from the literature (see, e.g., Refs. [4, 5]). To organize the appropriate information, one can classify these techniques by different criteria. For example, according to the abstraction levels of information processing [6, 7], they are grouped into spatial-domain methods [4, 5], methods of relevant features [8] and methods based on decision making [9, 10]. The vast majority of the available IF techniques are spatial-domain. In their turn, the spatial-domain techniques can be divided into pixel techniques that deal with single image pixels and super-pixel techniques that fuse the images on the level of pixel domains which include spatial-frequency domains. It is worth noting that most of the spatial-domain methods are designed for the target-data acquisition. The target-oriented IF techniques can be classified into *express methods* (employed for preliminary, real-time and urgent IF) and *office methods*. We argue that the distinction between the express and office methods is of crucial importance for several reasons. At a quick glance, it follows from their titles that this classification is inspired by the differences in the application areas and, consequently, the aims of the IF. However, the most evident argument for their distinction is that the requirements for the complexity of image-processing procedures and the image-quality assessment are essentially different for the two classes. In most cases, the requirements which are of primary importance for one of the classes are of no significance or even inapplicable for another class. Basing on the distinctive features of the express and office target-oriented methods (see the further analysis in this Section), we suggest a so-called *local* (*near-field*) *contrast* as a quantitative index used for the assessment of *visibility* of a target.

The office IF methods are well presented in the current literature. Their number increases permanently, their theoretical bases involve a number of modern scientific approaches, and their capabilities develop rapidly. On the other hand, only a few IF techniques available at the moment can be termed as the *express methods*. The main requirement for simplicity of the working algorithm implies that the express IF techniques should be based on simple mathematical or logical operations at the pixel level, which could be applied to the pairs of pixels of two fused images. In the field of real scalars, the number of principal algorithms is limited to a short list of basic arithmetic operations (addition, subtraction, multiplication and division) and some of their functional extensions (mainly exponentiation and extraction of roots). The most popular express IF technique is based on the algorithm of weighted addition (WA). Some other arithmetic operations have been used in different combinations. Most often, they are involved in complicated multistep algorithms [7] which cannot be considered as the express techniques.

Due to its simplicity, the WA-based IF technique is handy in processing and implies an easy interpretation of fusion results. For these reasons, it remains popular despite its essential drawback: the local contrast of a target in the WA-based IF is always lower than that of one of the partial

images [11]. For the fusion of Vis and infrared (IR) images, the situation is even worse. This is because, as a rule, the local contrasts of a target in the partial Vis and IR images have the opposite signs. Consequently, by its absolute value, the local contrast of the image fused with the WA technique might become lower than the contrasts of both partial images. Since the local contrasts of a target in the Vis and IR images are of the opposite signs, there might be a situation when the contrast of the target in the image fused with the WA method vanishes to zero, such that the target becomes invisible, even though it is well visible in the partial images [11]. Notice that, when the local contrast of the target vanishes, the integral contrast calculated for the image as a whole might not change so dramatically. The latter situation indicates that the integral contrast of the image, which is widely used for the assessment with the office methods, is not appropriate for the image-quality assessment of the target-oriented IF. Below we discuss the differences between the express and office target-oriented IF methods, revisit the approaches available in the literature for the calculation of contrast and consider them from the viewpoint of their applicability to the target-oriented IF, in particular to the WA method.

We claim that the problem of the local-contrast lowering (up to zero) which appears in the WA fusion can be overcome by passing from the field of real scalars to the fields of complex scalars or vectors. The principles of recently offered complex-scalar [12, 13] and complex-vector [14, 15] IF methods are presented in Section 2. Earlier (see Refs. [14, 15]) we have analyzed the images fused by the amplitude algorithm of a complex-vector IF (CVIF) method. In this work (see Section 3) we will perform the fusion using the angular algorithms of the same CVIF method for the same partial images (see Ref. [16]) and compare them with the images fused by some other algorithms. It is important that the angular algorithms of the CVIF method can provide a considerably higher quality of the fused images regarding the visibility of targets, if compared with the algorithms employed by the other express target-oriented methods. In the following discussion presented in this Section, we will scrutinize the assessment of quality of the images fused by the express target-oriented IF techniques, as opposed to those fused with the office IF methods. In Subsection 3.1, we will suggest a method for determining the local contrast of a target and relate this index to the visibility of this target. The results of quality comparison for the images fused by the angular-ellipticity algorithm of the CVIF method with the images obtained by some other express target-oriented methods will be presented in Subsection 3.2. Section 4 will show that, besides of a considerable enhancement of the target visibility provided by the complex-scalar IF and CVIF methods, they also satisfy all the requirements in order to be classified as the express target-oriented methods.

The office IF methods are designed, first of all, to enhance the overall quality of fused images. All the other characteristics of those methods (e.g., a processing time, a complexity of algorithms, a number of personnel involved and their knowledge of programming, a power of computers used and a cost of fused images) can be considered as secondary parameters subordinated to the main goal, a quality of images. Complicated multistep image-processing procedures involving image decomposition and transformation [17–21], deep learning [22–24], neural networks [25–27] and some other modern approaches are employed to enhance the quality of a fused image as a whole. The price is a considerably long processing time, a complexity of programming codes, a high power of computers, a highly qualified personnel of programmers performing the fusion and, as a consequence, a high overall cost of fused images. These drawbacks can be deemed as 'little inconveniences' whenever the IF is performed in a well-equipped office by the professionals inspired by the slogan '*Finis sanctificat media*'.

However, the IF should often be done in a real time using out-office applications. Moreover, there are many situations when the IF must be performed in the field or even battle conditions, using portative devices (smart cameras, smartphones, target sightseeing devices of military vehicles – tanks, boats, airplanes or space crafts – where the size of each additional device matters). Then the problem must be solved by an individual user or a crew having only beginner knowledge and skills in computer programming. Under these conditions, the 'little inconveniences' mentioned above would become unacceptable at all. Then the express IF techniques are needed for the out-office applications. The express IF can also be used as a preliminary fusion test before a more sophisticated method is applied. In most of the reviews on the present-day IF techniques, the WA-based IF method is referenced as a basic approach to which the quality of the images fused using the other modern methods is compared.

Another important argument for distinguishing the express target-oriented IF techniques from those used for the office applications is that the quantitative image-quality indices [6, 28] used for evaluating the images fused with the office methods are not applicable to the images obtained with the express methods. The reasons are multiple:

(i) The available quantitative image-quality indices are measured for the image as a whole and have an integral character which is expressed by an averaged (or relative) numerical value calculated over the whole image. This is exactly what is needed for the office methods, since they are developed to improve the quality of a whole imaged scene. On the contrary, the express target-oriented methods are aimed to enhance solely the visibility of targets, while the quality of the other areas, at least, does not matter. Most of the area located around a target is often not relevant to the visual search. At the least, improvement of irrelevant image areas is not needed. At the most, it might appear to be destructive for the very visual search, since enhancement of the visibility of irrelevant (i.e., not associated with a target) image details might lead to lateral masking [29–32] of this target, thus worsening its *conspicuity* [33].

(ii) The irrelevant areas of an image can appear totally extinguished, blurred or alighted (bleached) under the conditions of improper (too low or too high) lighting. Large improperly illuminated irrelevant areas are frequent for the Vis images taken at night or against intense-light sources (the sun, fire or artificial lighting). Such areas are present in almost all IR images. The presence of large improperly illuminated irrelevant areas worsens the overall quality of the image when it is assessed via the integral quality indices, though this fact does not necessarily affect the *visibility* and the *conspicuity* of a target. Sometimes this can even improve the latter characteristics.

(iii) The high quality indices calculated for an image as a whole do not guarantee a high visibility and conspicuity of a target. In most cases, the integral quality indices are just irrelevant to the visibility/conspicuity of the target. The entropy of the image, which represents a measure of its informativity, is a proper example. The higher the entropy of the image, the better is the informativity of the image as a whole [4]. However, the integral informativity of the image has little to do with the information about the target. Therefore, we are to conclude that the express target-oriented IF methods need their specific quality indices, which must be selectively associated with the target visibility and conspicuity.

Recently, new approaches have been offered for the express IF, which are based on expansion of the mathematical operations involved in the IF from the field of real scalars into the fields of complex scalars [12, 13] and complex vectors [14, 15]. The IF methods in the fields of complex scalars and vectors are analogues of the WA method in the field of real scalars. The

expansion to the fields of complex scalars and vectors provides a number of new algorithms applied on the pixel level to the pair of partial images. It enriches the information carried by the partial images, enhances their contrasts and, in particular, allows one to overcome the problem of lowering (or vanishing) contrast, which is known to affect considerably the images fused by the WA method in the real-scalar field [11]. What is of greater importance for the out-office applications, the scalar and vector IF techniques satisfy completely the requirements to the express target-oriented IF techniques. Below we will revisit the WA methods, as well as those related to the complex-scalar IF and CVIF.

As stressed above, the WA technique is the most popular in the express target-oriented IF. It is based on the fusion algorithm

$$\psi_w = w_u u + w_\upsilon \upsilon , \tag{1}$$

where w_u and w_υ are the weight coefficients for the partial images u and υ. At the condition $w_u = w_\upsilon = 1/2$, Eq. (1) reduces to the arithmetic-average algorithm given by

$$\psi_a = \frac{u + \upsilon}{2} . \tag{2}$$

At $w_u = w_\upsilon = 1$, it covers a simple-addition algorithm:

$$\psi_s = u + \upsilon . \tag{3}$$

The authors [13] argue that, while the WA algorithm (mutatis mutandis the simple-addition or arithmetic-average one) is well-applicable for fusing the mono-modal images, it is not appropriate for the multimodal images for several reasons. To prove the latter statement, we have to consider a notion of local contrast of a target.

As remarked by Westheimer [34], there is no standard definition for the contrast. The contrast parameters introduced by Weber and Michelson are most commonly used in the literature. The first parameter is given by

$$c_W = \frac{I_t - I_b}{I_b} , \tag{4}$$

where I implies the intensity (brightness) and the indices t and b correspond to a target and a background, respectively. The c_w values fall in the range $[-1, +\infty[$. The Michelson contrast is defined via the maximal ($I_{\max}$) and minimal ($I_{\min}$) intensities as

$$C_M = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}} , \tag{5}$$

and varies in the range $[0, +1]$. This contrast parameter has been introduced to describe the contrast of a sinusoidal periodic pattern with the values $I_{\max}$ and $I_{\min}$ corresponding to the extreme (maximal and minimal) brightness values of a pattern. Another version of the Michelson contrast,

$$c_M = \frac{I_{\max} - I_{\min}}{\overline{I}} = 2\frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}} = 2C_M , \tag{6}$$

varies in the range $[0, +2]$, with $\overline{I} = \left(I_{\max} + I_{\min}\right)/2$ standing for the arithmetic average of the extrema. Note that the values of the Michelson contrast are always positive, which is inconvenient regarding the Vis and IR images in which a target can be, respectively, 'positive' and 'negative' in the conventional sense of positive and negative images in the image-processing theory.

It is also worth noting that the value ranges for the Weber (see Eq. (4)) and Michelson (see Eq. (5)) contrasts are not symmetric with respect to zero, which is rather inconvenient. To overcome these inconveniences, we adopt the notion of a local contrast defined in Ref. [11] as

$$k = \frac{I_b - I_t}{\bar{I}}, \qquad (7)$$

where $\bar{I} = (I_b + I_t)/2$ is the average brightness of the background (b) and the target (t). The values of the local contrast given by Eq. (7) fall in the range [–2, 2]. Although the parameter given by Eq. (7) is similar to the Michelson contrast defined by Eq. (6), the range of its changes is symmetric with respect to zero. Indeed, one can have $I_t < I_b$ (a positive contrast), $I_t = I_b$ (a zero contrast) or $I_t > I_b$ (a negative contrast) in case of the local contrast, whereas for the Michelson contrast one has always the relation $I_{\max} > I_{\min}$. For this reason, the notion of the local contrast is more convenient for the target-oriented image-quality assessment, if compared with the Michelson and Weber contrasts.

The procedures used for experimental measurements of the local contrast will be proposed in Subsection 3.1. The normalized values of the local contrast varying in the range [–1, 1] are obtained by inserting the factor of $1/2$, similarly to the relation between Eqs. (5) and (6) for the Michelson contrast:

$$K = \frac{1}{2}k = \frac{I_b - I_t}{I_b + I_t}. \qquad (8)$$

A following remark is worth mentioning here. Let us notice that the definitions of the local contrast introduced by Eqs. (7) and (8) differ by their signs from a so-called 'target contrast' introduced in the other works. For example, in Ref. [35] it is accepted that the positive target contrast corresponds to the situation when the brightness of the target is higher than that of its background and, vice versa, a target which is seen as a dark object against a bright background has a negative contrast. We argue that such a definition contradicts the concepts of positive and negative photography traditionally adopted in the image-processing theory. Usually, a target absorbs the Vis light and, consequently, it looks like a dark object against a bright background in any image taken in the Vis light (see, e.g., a Vis image presented in Fig. 1a below). Such a Vis image corresponds to a natural look of the target as observed with naked eyes and, quite reasonably, the very Vis image is called a positive image, in contrast to a negative image where the brightness order is reversed, at least for the target.

It is also worth recalling that a conventional negative image is obtained on a film through a chemically developing photography. In this respect, an IR image should also be considered a negative image. Usually, a target emits or reflects the IR light and, consequently, it is seen in an IR image as a bright object against a dark background. For this reason, we follow the traditional concept of the positive and negative photography when defining the local contrast. Hence, the positive sign in Eqs. (7) and (8) corresponds to a dark target seen against a bright background and, vice versa, the negative sign of the local contrast corresponds to a bright target observed against a dark background. Of course, it is not a rigid rule that a target in an IR image should look bright against a dark background: a gun and a knife in the IR images displayed in Fig. 1b and Fig. 2b are examples. Note that the sign of the target contrast might be of no importance at the stage of detection of the target. Usually, the silhouette of a person or the contour of an object is enough for

detection. However, the sign of the local contrast is of crucial importance when we recognize and identify the target.

With respect to the simple addition (see Eq. (3)), the local contrast k_s of the target in the fused image is a linear function of the local contrasts k_u and k_υ of the target in the partial images [11]:

$$k_s = \omega_u k_u + \omega_\upsilon k_\upsilon, \tag{9}$$

where

$$\omega_u = \frac{\overline{u}}{\overline{u} + \overline{\upsilon}}, \omega_\upsilon = \frac{\overline{\upsilon}}{\overline{u} + \overline{\upsilon}}, k_u = \frac{\Delta u}{\overline{u} + \overline{\upsilon}},$$
$$k_\upsilon = \frac{\Delta \upsilon}{\overline{u} + \overline{\upsilon}}, \overline{u} = \frac{u_b + u_t}{2}, \overline{\upsilon} = \frac{\upsilon_b + \upsilon_t}{2}, \tag{10}$$
$$\Delta u = u_b - u_t,$$

with u_t, u_b and υ_t, υ_b being the brightness values of the target (t) and the background (b) in the Vis (u) and IR (υ) images, respectively. It follows from Eq. (9) that, because of the opposite signs of the local contrasts k_u and k_υ of the target in the partial Vis (u) and IR (υ) images, the local contrast k_s of the target in the image fused by the simple-addition algorithm (see Eq. (3)) is always lower than those in the partial images, such that one has $k_s = 0$ for $\omega_u k_u = -\omega_\upsilon k_\upsilon$. Indeed, substitution of Eqs. (10) in Eq. (9) gives

$$k_s = \frac{\Delta u + \Delta \upsilon}{\overline{u} + \overline{\upsilon}}. \tag{11}$$

It follows from Eq. (11) that the local contrast k_s of the target in the fused image is zero whenever we have $\Delta u = -\Delta \upsilon$. In case of the weighted-addition algorithm given by Eq. (1), the local contrast of the fused image acquires the form similar to that for the simple-addition algorithm [11]:

$$k_w = \omega_{uw} k_u + \omega_{\upsilon w} k_\upsilon, \tag{12}$$

with

$$\omega_{uw} = \frac{w_u \overline{u}}{w_u \overline{u} + w_\upsilon \overline{\upsilon}},$$
$$\omega_{\upsilon w} = \frac{w_u \overline{\upsilon}}{w_u \overline{u} + w_\upsilon \overline{\upsilon}}. \tag{13}$$

It is clear that Eqs. (12) and (13) for the WA algorithm reduce to Eqs. (9) and (10) for the simple-addition algorithm at $w_u = w_\upsilon = 1$. Similarly to the simple-addition algorithm, the local contrast of a target in the image fused by the WA algorithm is lower than the local contrasts of the partial Vis and IR images; it becomes zero at the condition $\omega_{uw} k_u = -\omega_{\upsilon w} k_\upsilon$.

Concluding this Subsection, we state that addition of (in general, multimodal) Vis and IR images in the field of real scalars is at least unjustified for the following reasons.

1. Vis and IR images differ in the physical principles of their detection. Although they are stored in a computer memory as tables of brightness values measured in the same units, there is no evidence that these brightnesses correlate with each other. In other words, they are independent and one cannot transform one into another by an evident transformation. This situation is akin to numerous cases in physics, such as linear and circular birefringences or refractive and absorption indices in optics, real and imaginary parts of a complex viscoelastic modulus in hydrodynamics, a complex frequency and a complex resistance in electronics, etc. They cannot be added in the field of real scalars but form complex physical quantities.

68

Ukr. J. Phys. Opt. 2023, Volume 24, Issue 1

2. A target usually looks dark against a bright background in a Vis image (a positive local contrast). Unlike this situation, a target in an IR image looks bright against a dark background (a negative local contrast). The latter is a consequence of so-called Prevost rules [36], according to one of which the better a body absorbs the electromagnetic irradiation the better it irradiates it. This qualitative rule is expressed quantitatively by the Kirchhoff law [37], which states that the irradiation capability of a body is proportional to its absorption capability at the same temperature. As a result, the objects which look dark in the Vis images in many cases look bright in the IR images.

We note that the Prevost rule and the Kirchhoff law apply to a body that stays in a thermal equilibrium, with no internal energy sources present. In addition to the equilibrium (passive) thermo-radiation of the bodies and the engines of machines, which represent typical targets in a military sense, these objects have internal sources of energy and thus irradiate IR waves much intensely than the environment in the equilibrium. The very possibility of the active irradiation of IR waves due to some internal energy sources would enhance the negative contrast of a target in an IR image. Therefore, one is led to conclude that the local contrasts of the targets in Vis and IR images are often of the opposite signs. However, the statement that the contrasts of a target in these images differ by their signs does not mean that, by inversion of the contrast, a negative IR image will be transformed into a positive image similar to a Vis-image. Although the relationships between the IR and Vis images of the same scene involves the Prevost rules and the Kirchhoff law, the appearance of a real IR image is complicated by interplay between the absorption, the reflection and the emission of IR waves.

Eq. (12) shows that the local contrast of a target is lowered considerably when a Vis image with the positive local contrast of a target is fused by the WA algorithm with an IR image showing the negative local contrast for the same target in the same scene. Lowering (down to zero) of the local contrast observed with the simple and WA algorithms [11] is another argument against the addition of Vis and IR images within the field of real scalars. Extension into the field of complex numbers is one of the reasonable possibilities to solve this problem.

3. If multiple Vis and IR images are available, one faces an uncertainty in the formation of image pairs for fusion. A reasonable approach would be to fuse separately the Vis and IR images into one pair of pre-fused Vis and IR images and then to fuse these two pre-fused images. This operation resembles a mathematical operation of addition of complex numbers, in which their real and imaginary parts add separately.

At least the three arguments presented above prove that a fusion of Vis and IR images must be done in the field of complex numbers rather than in the field of real scalars.

2. Complex IF methods
2.1. Complex-scalar IF method
In the framework of the complex-scalar IF method [12, 13], one of the two multimodal images (a Vis image u or an IR image υ) can be chosen as the real part and the other one (an IR image υ or a Vis image u) as the imaginary part of a complex scalar function ψ. Generally, there are no restrictions on the choice of the real and imaginary parts of ψ. Thus, the corresponding relation can be written in the two forms:

$$\psi_{neg} = \frac{1}{\sqrt{2}}(u + i\upsilon) = |\psi| e^{i\varphi_{neg}} , \tag{14}$$

$$\psi_{pos} = \frac{1}{\sqrt{2}}\left(\upsilon + iu\right) = \left|\psi\right| e^{i\varphi_{pos}}\,, \tag{15}$$

where

$$\left|\psi\right| = \left|\psi_{neg}\right| = \left|\psi_{pos}\right| = \frac{1}{\sqrt{2}}\sqrt{u^2 + \upsilon^2} \tag{16}$$

is the amplitude and

$$\varphi_{neg} = \arctan\left(\upsilon/u\right), \tag{17}$$

$$\varphi_{pos} = \arctan\left(u/\upsilon\right) \tag{18}$$

are the phases of the complex functions ψ_{neg} and ψ_{pos}, respectively. The factor $1/\sqrt{2}$ in Eqs. (14)–(16) implies normalization of the amplitude. Namely, one has the maximal amplitude $\left|\psi_{\max}\right| = 1$ from Eq. (16) at the maximal brightness values $u_{\max} = 1$ and $\upsilon_{\max} = 1$.

Eqs. (16)–(18) can be used as working algorithms for fusion of the input images u and υ. Namely, we have the amplitude algorithm given by Eq. (16), which is common for the two forms ψ_{neg} and ψ_{pos}, and the two phase algorithms φ_{neg} and φ_{pos} given by Eqs. (17) and (18). The latter can also be rewritten as

$$t_{neg} = \tan\varphi_{neg} = \upsilon/u\,, \tag{19}$$

$$t_{pos} = \tan\varphi_{pos} = u/\upsilon\,. \tag{20}$$

The forms given by Eqs. (17)–(20) for the phase φ- and t-algorithms imply a mathematic singularity, i.e. a divergence when the denominators tend to zero. Since the partial Vis (u) and IR (υ) images are usually taken at low lighting, the zeros of u and υ are quite plausible. To avoid the singularities, one can add a small parameter ε to the denominators and Eqs. (17)–(20). Then their modified forms become as follows:

$$\varphi_{neg}^{\varepsilon} = \arctan\frac{\upsilon}{u+\varepsilon}\,, \tag{21}$$

$$\varphi_{pos} = \arctan\left[\frac{u}{\upsilon+\varepsilon}\right]\,, \tag{22}$$

$$t_{neg} = \tan\varphi_{neg}^{\varepsilon} = \frac{\upsilon}{u+\varepsilon}\,, \tag{23}$$

$$t_{pos} = \tan\varphi_{pos}^{\varepsilon} = \frac{u}{\upsilon+\varepsilon}\,. \tag{24}$$

The two types of the algorithm (i.e., the amplitude and phase ones) carry the independent complementary information. The amplitude algorithms are based on the binary quadratic forms constructed of the two partial images, so that they carry the compositional information on targets (their numbers and location). The phase algorithms are based on the operation of division and so carry the relative information on the capability of objects to emit/reflect the IR waves, thereby marking potential targets.

2.2. Complex-vector function algorithms

2.2.1. Amplitude, azimuth and ellipticity in the eigen coordinate system. In the framework of the CVIF method [14, 15], one of the two multimodal (Vis u or IR υ) images is chosen as the real component and the other one (IR υ or Vis u) as the imaginary component of the complex vector $\vec{\psi}^{(0)}$. Similarly to the complex-scalar IF method [12, 13], there are no restrictions on the choice of

the real and imaginary components of $\vec{\psi}^{(0)}$. Hence, it can be written in the two alternative forms:

$$\vec{\psi}_{neg}^{(0)} = \frac{1}{\sqrt{2}} \begin{bmatrix} u \\ i\upsilon \end{bmatrix}, \tag{25}$$

$$\vec{\psi}_{pos}^{(0)} = \frac{1}{\sqrt{2}} \begin{bmatrix} \upsilon \\ iu \end{bmatrix}. \tag{26}$$

Eqs. (25) and (26) are symbolic representations of the complex vectors $\vec{\psi}_{neg}^{(0)}$ and $\vec{\psi}_{pos}^{(0)}$. Their explicit representations are as follows:

$$\vec{\psi}_{neg}^{(0)} = \frac{1}{\sqrt{2}} \left(u\vec{c}_1 + i\upsilon\vec{c}_2 \right), \tag{27}$$

$$\vec{\psi}_{pos}^{(0)} = \frac{1}{\sqrt{2}} \left(\upsilon\vec{c}_1 + iu\vec{c}_2 \right), \tag{28}$$

where $\vec{c}_1$ and $\vec{c}_2$ are the orthogonal unit vectors along which the axes of the Cartesian coordinate system are directed. Hereafter, this coordinate system is called an 'eigen coordinate system'. By their forms, the complex vectors $\vec{\psi}_{neg}^{(0)}$ and $\vec{\psi}_{pos}^{(0)}$ bear a striking resemblance to the Jones vector describing the electric field of an elliptically polarized light wave [38–41], with the partial (Vis u and IR υ) images corresponding to the vector components of the light electric field. Basing on this analogy, one can introduce the amplitudes $\left| \vec{\psi}_{neg,pos}^{(0)} \right|$, the azimuth angles $\chi_{neg,pos}^{(0)}$ and the ellipticities $\gamma_{neg,pos}^{(0)}$ of the vectors $\vec{\psi}_{neg}^{(0)}$ and $\vec{\psi}_{pos}^{(0)}$. In the eigen coordinate system of the vector basis, we find

$$\left| \vec{\psi}^{(0)} \right| = \left| \vec{\psi}_{neg}^{(0)} \right| = \left| \vec{\psi}_{pos}^{(0)} \right| = \frac{1}{\sqrt{2}} \sqrt{u^2 + \upsilon^2}, \tag{29}$$

$$\tan 2\chi_{neg,pos}^{(0)} = 0, \tag{30}$$

$$\sin 2\gamma_{neg,pos}^{(0)} = 2\frac{u\upsilon}{u^2 + \upsilon^2}. \tag{31}$$

It is worth noting that Eq. (29) for the complex-vector amplitude $\left| \vec{\psi}^{(0)} \right|$ is the same for the vectors $\vec{\psi}_{neg}^{(0)}$ and $\vec{\psi}_{pos}^{(0)}$ defined in their eigen coordinate system. It coincides with Eq. (16) for the amplitude $|\psi|$ of the scalar function given by Eqs. (14) and (15). It follows from Eq. (30) that the azimuth angle becomes zero ($\chi_{neg,pos}^{(0)} = 0$), as it should be for the eigen coordinate system. Eq. (31) for the ellipticity angle represents a new algorithm for the fusion of the partial images u and υ. Application of the amplitude algorithm of the CVIF method has been introduced in Ref. [14]. In Subsection 3.2, we will present for the first time the IF based upon the ellipticity algorithm given by Eq. (31).

2.2.2. Amplitude, azimuth and ellipticity under Jones-matrix transformation. Using the light-wave analogy in the framework of the Jones formalism [14, 15], one can transform the vectors $\vec{\psi}_{neg}^{(0)}$ and $\vec{\psi}_{pos}^{(0)}$ to

$$\vec{\psi}_{neg,pos} = J\psi_{neg,pos}^{(0)}, \tag{32}$$

where J is a 2×2 matrix of complex elements, an analogue of the Jones matrix describing the optical properties of an optically anisotropic medium through which a light wave propagates.

2.2.3. Transformed amplitude algorithm. The amplitude of the transformed vector $\vec{\psi}_{neg,pos}$ given by Eq. (32) acquires the form

$$\left|\vec{\psi}\right| = \sqrt{(au)^2 + (bv)^2 + c^2 uv} \,, \tag{33}$$

where the parameters a, b, and c are the quadratic forms of the elements of the matrix J. Their explicit forms have been described in Ref. [14]. The indices *neg* and *pos* in Eq. (33) are omitted since, in fact, the parameters a, b and c play the roles of weight coefficients of the quadratic terms u^2, v^2 and uv in the binary quadratic form composed of the partial images u and v [14]. Due to the normalization condition, the vector amplitude $\left|\vec{\psi}\right|$ varies in the range $[0;1]$. Then the weight coefficients vary in the same range and obey the normalization rule

$$a^2 + b^2 + c^2 = 1 \,. \tag{34}$$

A number of particular cases for the parameters a, b and c are worth considering:

$\left|\vec{\psi}\right|\big|_{a=1,b=c=0} = u$ is the partial Vis image;

$\left|\vec{\psi}\right|\big|_{b=1,a=c=0} = v$ is the partial IR image;

$\left|\vec{\psi}\right|\big|_{a=b=1/\sqrt{2},c=0} = |\psi| = \left|\vec{\psi}^{(0)}\right| = \left|\vec{\psi}_{neg}^{(0)}\right| = \left|\vec{\psi}_{pos}^{(0)}\right| = \dfrac{1}{\sqrt{2}}\sqrt{u^2 + v^2}$ is the root mean square (RMS)

for the partial Vis and IR images, which coincides with the amplitude algorithms in the field of complex scalars (see Eq. (16)) and complex vectors (see Eq. (29));

$\left|\vec{\psi}\right|\big|_{c^2=2ab} = au + bv$ is the weighted-addition algorithm in the field of real scalars for the partial Vis and IR images (see Eq. (1));

$\left|\vec{\psi}\right|\big|_{a=b=c^2=1/2} = \psi_a = \dfrac{u+v}{2}$ is the arithmetic mean for the partial Vis and IR images (see Eq. (2));

the relation

$$\left|\vec{\psi}\right|\big|_{c=1,a=b=0} = \sqrt{uv} \,, \tag{35}$$

is the geometric mean of the partial Vis and IR images.

By varying the weight coefficients in the range $[0;1]$, one can use Eq. (33) together with the normalization condition given by Eq. (34) as an algorithm for the dynamic IF. Then, instead of a single image fused from two partial images, one synthesizes a fused video [14, 15]. Playing this video, an operator can stop it at the moment of the best visibility/conspicuity of a target. Note that the video format of the IF enhances the conspicuity of the target in comparison with the observation of single images.

2.2.4. Transformed azimuth and ellipticity algorithms. The azimuth and the ellipticity of the vector $\vec{\psi}_{neg,pos}$ transformed by Eq. (32) read as [14]

$$\tan 2\chi_{neg,pos} = \left.\frac{\xi_{\chi n}^2 u^2 + \mu_{\chi n}^2 v^2 + \sigma_{\chi n}^2 uv}{\xi_{\chi d}^2 u^2 + \mu_{\chi d}^2 v^2 + \sigma_{\chi d}^2 uv}\right|_{neg,pos} , \tag{36}$$

$$\sin 2\gamma_{neg,pos} = \left.\frac{\xi_{\gamma n}^2 u^2 + \mu_{\gamma n}^2 v^2 + \sigma_{\gamma n}^2 uv}{\xi_{\gamma d}^2 u^2 + \mu_{\gamma d}^2 v^2 + \sigma_{\gamma d}^2 uv}\right|_{neg,pos} , \tag{37}$$

where the coefficients associated with the quadratic terms u^2, υ^2 and $u\upsilon$ in Eqs. (36) and (37) are composed of the components of the matrix J (the explicit forms of the latter can be found Ref. [14]). It is seen from Eqs. (36) and (37) that the angular algorithms for the azimuth and ellipticity have similar forms which differ only by the explicit relations for the coefficients of the quadratic terms. For this reason, below we will focus only on the ellipticity algorithm given by Eq. (37).

Similarly to the complex-scalar IF method, the CVIF method provides the two types of fusion algorithms. These are the amplitude and angular algorithms, which carry the independent complementary information. The amplitude algorithms are based on the binary quadratic forms constructed of two partial images and they carry the compositional information on targets (their numbers and location). The azimuth and ellipticity algorithms (see Eqs. (36) and (37)) are based on the operation of division and so they carry the relative information on the capability of objects to emit/reflect IR waves. This can mark potential targets.

The next Section is devoted to applications of the complex-scalar IF and the CVIF algorithms to realistic pairs of the partial Vis and IR images. In particular, in Subsection 3.1 we introduce a quantitative index characterizing the quality of target visibility and present a method for determining the local contrast of a target. After that, we assess the partial and fused images in Subsection 3.2.

3. Experimental results and discussion
3.1. Method for determining the local contrast of targets
Our method differs from its analogues by the measurement procedure: the contrast measured by us characterizes the very target but not the image as a whole. To calculate the integral contrast for an image, commonly one divides the image into blocks, inside of which the contrast of a central (target) part is calculated with respect to a periphery (background), and then the contrast values for different blocks are averaged [7]. The number of blocks is chosen from tens to hundreds, depending on the complexity of the image scene. It is clear that, strictly speaking, the averaged image contrast calculated in the above manner does not specify the contrast of any target in the image.

To calculate the local contrast of a target, we suggest measuring the brightness profile (see, e.g., the plots above Figs. 1a and 1b) along a horizontal line (see Figs. 1a and 1b) drawn across this target (e.g., faces of left and right persons or a gun and a knife in Figs. 1–3). The probing line, along which the brightness is measured, is chosen to be horizontal in view of the fact that a visual search is conventionally performed along horizontal lines/stripes. The procedure of calculation of the local contrast from the data of the brightness profile is illustrated in Fig. 1, where the horizontal lines are drawn across the face of the right person in the Vis and IR images at the same vertical coordinate. The normalized local contrast shown in Fig. 1c is calculated using Eq. (8) where the brightness data measured along the line is used as I_t (see the data plotted above the Vis and IR images in Figs. 1a and 1b) and the I_b value is a constant brightness measured for the background in the vicinity of the target.

It is worthwhile that the target backgrounds in Figs. 1a and 1b are practically uniform and, thus, there is no difficulty to determine the local value I_b of the background brightness. However, the brightness of the background might vary considerably across the image in a general case. A definition of the local band-limited contrast has been given in Ref. [42]: a contrast value is

assigned to every point of the image as a function of a spatial-frequency band. With such a definition, all the points in the image are of the same importance, which means that the image-quality assessment is not target-oriented. In a target-oriented assessment, a researcher is interested first of all in the image quality of targets. Lowering of the image quality in all the other non-target (i.e., non-relevant) points makes the background more uniform and thus might be even desirable, since it enhances the conspicuity of the target.
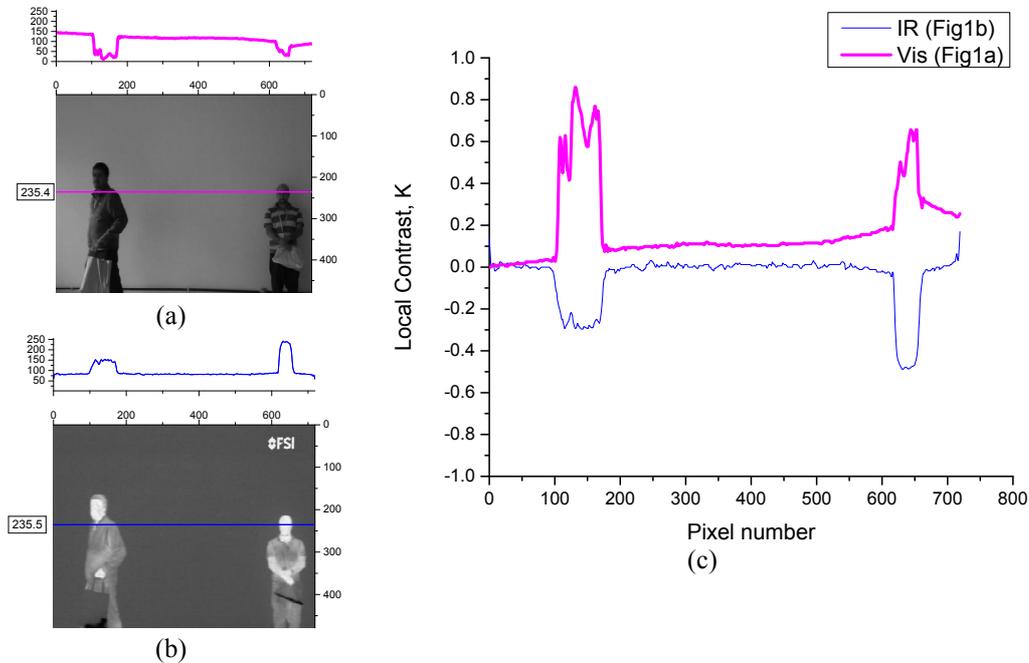
Fig. 1. (Colour online). Illustration of determination of the local contrast of a target. Brightness profiles measured along horizontal lines drawn across left and right persons at the same vertical coordinate (≈ 235) on the partial Vis (a) and IR (b) images are plotted as thick and thin solid lines above the images. Normalized local contrasts K for the left and right persons are calculated using Eq. (8) and the brightness profiles measured from the partial Vis image shown in Fig. 1a (a thick line), the IR image shown in Fig. 1b (a thin line).

In case of a non-uniform background in the target-oriented image-quality assessment, we propose to calculate I_b as the brightness average over a line segment of the length equal, say, to the half-width of the target along the line in the immediate vicinity of the target. A possibility for graphical representation of the local contrast is an advantage of our method. It allows to depict the visibility of a target as a function $K(x)$ of the horizontal coordinate x (see, e.g., Fig. 1c). On the contrary, the traditional approaches compute the image contrast using a programming code and display it as a single parameter. Conventionally, the visibility of a target represents the contrast of the latter divided by the threshold value of the contrast K_{th} which is still resolved by human eyes. Therefore, the higher the band corresponding to the target in the plot $K(x)$, the higher the target visibility is.

Since the background brightness I_b is measured in the immediate vicinity of the target, the target is always visualized in the shape of a step-like band in the plot $K(x)$, unless we deal with a vanishingly low contrast of the target when the target is hardly visible. Since the $I_t(x)$ value is chosen as a function of the horizontal coordinate x, there might be multiple step-like bands in the $K(x)$ plot, which correspond to some other objects in the image. For instance, one can find two

such step-like bands in Fig. 1c, which correspond to different (left and right) persons in the images shown in Figs. 1a and 1b. Were more persons or other objects present in the images, still more step-like bands would have been observed in the $K(x)$ plot. Multiple step-like bands corresponding to the objects different from the target in the image do not affect the visibility of this target but signal us about a lowered conspicuity of the target. Multiple objects, which are similar to the target and stand around it, would mask the target and so lower its conspicuity. Here we do not explore possible relations between the visibility and conspicuity of the target. A method for determining the conspicuity of the target from the target-brightness profile will be a subject of our forthcoming work.

The local-contrast profile shown in Fig. 1c indicates clearly the positions of the left and right persons in the Vis (a thick pink solid line) and IR (a thin blue solid line) partial images. These positions correspond to the step-like bands limited by the two almost vertical bounding line segments. The contrast obtained by averaging the values in between these bounding segments is taken as the normalized local contrast of the target (see Table 1).

It is seen that the local-contrast values for the target deduced from the Vis and IR images have indeed the opposite signs: the sign is positive for the positive (in the conventional sense of positive photography) Vis image and negative for the apparently negative IR image. By saying 'apparently negative' we imply that the IR image is not truly a negative image as a whole: only the targets which emit the IR waves appear to be negative in the IR image.

The above measurements and calculations of the normalized local contrast have been performed for all of the four targets (the faces of the left and right person, the gun and the knife) in the input partial Vis and IR images (Fig. 2). The same has been done for the images fused by the arithmetic-average algorithm (Fig. 3a) in the field of real scalars and by the RMS (Fig. 3b), geometric-mean (Fig. 3c) and ellipticity (Fig. 3d) algorithms in the field of complex vectors. The results are gathered in Table 1.

3.2. Assessment of partial and fused images

The partial (Figs. 2a, b) and fused (Fig. 3) images taken from Ref. [16] include two men holding bags in their hands. One can recognize the faces of both men in the partial Vis image shown in Fig. 1a. However, it is impossible to see whether they carry any hidden dangerous objects. This observation agrees with the local-contrast values of the targets presented in Table 1. Indeed, relatively high positive values $K = 0.65$ and 0.54 measured respectively for the faces of the right- and left persons in the partial Vis image of Fig. 1a quantitatively reflect a naturally positive brightness observed in the Vis image in the traditional sense of positive and negative photography, which is required for the face recognition.

On the contrary, the corresponding negative local-contrast values $K = -0.49$ and -0.47 reveal that the IR images of the faces of these persons are far from their real natural look. Although this conclusion seems to be trivial for an operator inspecting the images visually, it is of crucial importance for the computer-based target recognition. Note that, in general, a target-containing image can be considered to have a sufficient quality if it provides high enough absolute values and appropriate signs of the local contrasts simultaneously for all targets involved. Namely, the local contrast must be positive for the faces of persons and it can have any sign for dangerous objects as long as it is high enough by its absolute value in order to detect and recognize these objects by their shape. A naturally positive contrast of the dangerous objects might be needed at the stage of their identification.

Table 1. Objective target-oriented and integral quality assessment for partial and fused images.

Image	Normalized local contrast, K				Integral objective metric	
	Left person		Right person			
	Face	Gun	Face	Knife	E	SD
Vis	0.65	0	0.54	0	4.27	0.12
IR	−0.49	0.71	−0.47	0.36	3.02	0.10
Arithmetic average	−0.17	0.31	−0.25	0.21	4.32	0.05
RMS	−0.17	0.4	−0.19	0.24	6.67	0.05
Geometric mean	0.31	0.58	0.20	0.26	6.56	0.06
Ellipticity	0.70	0.60	0.55	0.15	6.66	0.19

Although negative local-contrast values measured for the faces of persons signal that the IR image is not applicable for the face recognition, it allows one to see that the men on the left and right sides have in their bags the objects resembling respectively a gun and a knife, which cannot be seen in the partial Vis image. These conclusions based on the visual inspection correlate with the quantitative data given in Table 1. The local contrasts for the gun and the knife in the IR image are respectively equal to $K = 0.71$ and 0.36, whereas the corresponding parameters measured from the Vis image are zero as they should be, since these objects are not visible in the Vis image. We conclude therefore that the notion of the local contrast is an appropriate quantitative image-quality index, which is suitable for the computer detection, recognition and, possibly, identification of targets.

In this work, the IF is employed to synthesize an image suitable for a simultaneous detection of dangerous objects and face recognition. In such a case, the IF is definitely target-oriented. The faces of the persons and the dangerous objects (the gun and the knife) are the targets of the IF.

(a) (b)

Fig. 2. Vis (a) and IR (b) input images.

The image shown in Fig. 3a is fused in the field of real scalars by the arithmetic-average algorithm (see Eq. (2)). On the other hand, the image of Fig. 3b is fused in the field of complex vectors using the RMS algorithm (see Eq. (29)), which is equivalent to the amplitude algorithm in the field of complex scalars (see Eq. (16)). A visual inspection of Figs. 3a and 3b reveals that these algorithms do not provide a proper (and simultaneous) appearance of the faces of persons and the dangerous objects carried in their bags: although the gun and the knife are visible, the faces of the

persons cannot be recognized. The latter conclusion is supported by the measured local contrasts, which have the negative signs for the faces. Namely, we have obtained $K = -0.17$ for the left person in both images fused by the arithmetic-average and RMS algorithms. The same parameters are equal to $K = -0.25$ and -0.19 for the right person in the images fused by the arithmetic-average and RMS algorithms, respectively.
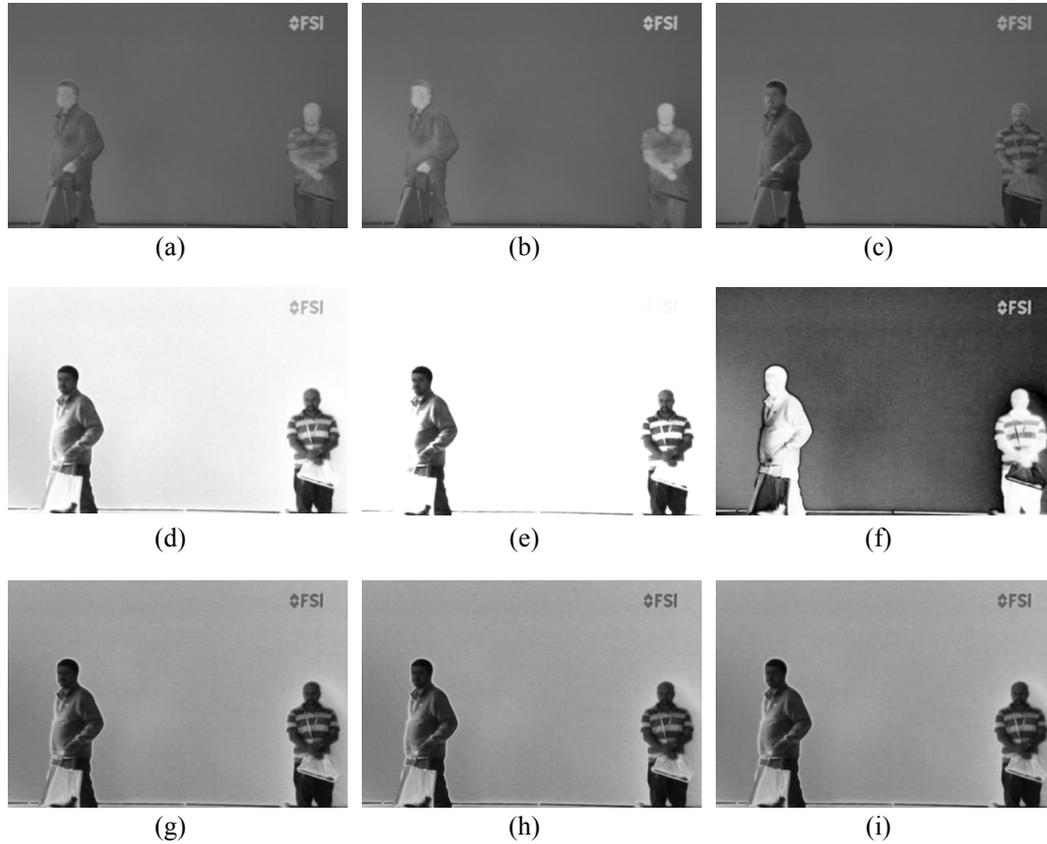
Fig. 3. Images fused by different algorithms: (a) arithmetic average, (b) RMS (see Eq. (29)), (c) geometric mean (see Eq. (35)), (d) ellipticity $\sin 2\gamma_{neg,pos}^{(0)} = 2u\upsilon / \left(u^2 + \upsilon^2 \right)$ (see Eq. (31)), (e) $\tan\left(2\gamma_{neg,pos}^{(0)} \right) = 2u\upsilon / \left| u^2 - \upsilon^2 \right|$, (f) $\cos\left(2\gamma_{neg,pos}^{(0)} \right) = \left| u^2 - \upsilon^2 \right| / \left(u^2 + \upsilon^2 \right)$, (g) ellipticity-angle calculated as $\gamma_{neg,pos}^{(0)} = (1/2)\arcsin\left(2u\upsilon / \left(u^2 + \upsilon^2 \right) \right)$ (calculated based on Eq. (31)), (h) $\gamma_{neg,pos}^{(0)} = \arctan\left(2u\upsilon / \left| u^2 - \upsilon^2 \right| \right)$, and (i) $\gamma_{neg,pos}^{(0)} = \arccos\left(\left| u^2 - \upsilon^2 \right| / \left(u^2 + \upsilon^2 \right) \right)$.

The contrary situation occurs with the image of Fig. 3c, which is fused by the geometric-mean algorithm (see Eq. (35)) representing a particular case of the amplitude algorithm in the field of complex vectors. Here one can recognize the faces of both persons and see that they carry dangerous objects in their bags, which resemble a gun and knife. The positive local contrasts $K = 0.31$ and 0.20 measured respectively for the faces of the left and right persons indicate that the image fused by the geometric-mean algorithm is suitable for the computer face recognition. Still higher quality is typical for the image of Fig. 3d, which is fused by the ellipticity algorithm (see Eq. (31)) in the field of complex vectors. Indeed, the corresponding values $K = 0.70$ and 0.55 of the local contrast are positive and much higher than those obtained for the geometric-mean algorithm.

We have fused the partial images with the other trigonometric algorithms of the ellipticity angle. This has been done using the relation

$$\tan\left(2\gamma_{neg,pos}^{(0)}\right) = 2\frac{u\upsilon}{\left|u^2 - \upsilon^2\right|} \tag{38}$$

(see Fig. 3e), the relation

$$\cos\left(2\gamma_{neg,pos}^{(0)}\right) = \frac{\left|u^2 - \upsilon^2\right|}{u^2 + \upsilon^2} \tag{39}$$

(see Fig. 3f) and the algorithm of the ellipticity angle $\gamma_{neg,pos}^{(0)}$ defined by Eq. (31) (see Fig. 3g).

To verify the adequacy of our calculations, we have also fused the partial images using the ellipticity-angle algorithms deduced from Eq. (38) (see Fig. 3h) and Eq. (39) (see Fig. 3i). As expected, all of the three images fused by different ellipticity-angle algorithms obtained from different trigonometric functions, Eqs. (31), (38) and (39), appear to be equivalent (compare the data of Figs. 3g–i with each other).

Fig. 4. Dynamic IF based on the ellipticity algorithm given by Eq. (40) parameterized by $a \in [-1;1]$ and $b \in [1;-1]$. To play video in the on-line version of article, click on the image.

It is worth noting that the local-contrast values measured by us (see Table 1) correlate well with the visibility of the targets and the possibilities of their recognition in the images (see Figs. 2 and 3). On the contrary, the integral image-quality indices such as the image entropy (E) and the standard deviation (SD) measured for the same images (see also Table 1) hardly correlate with the image quality of the targets. The image-entropy values $E = 6.67$, 6.56 and 6.66 measured for the images fused respectively by the RMS, geometric-mean and ellipticity algorithms are considerably higher than those obtained for the input Vis and IR images and the image fused by the arithmetic-mean algorithm (respectively $E = 4.27$, 3.02 and 4.32). Although higher entropies might imply a higher overall quality of the RMS-, geometric-mean- and ellipticity-based images, the entropy values do not indicate explicitly which of the images can be better used for the face recognition and the detection of dangerous objects. A similar conclusion follows from the SD data (see Table 1). Although the SD data might carry some information on the overall quality of the images, they are useless in drawing the conclusions about the possibilities for successful face recognition from the images and the detection of dangerous objects. Therefore the entropy and SD figures representing the integral indices cannot be applied to the target-data acquisition, which differs drastically from the local-contrast figures derived by us. It is understood that this conclusion would remain valid with respect to the other quantitative image-quality indices introduced for the IF and calculated as integral parameters over the whole image. The integral image-quality indices such as

the integral contrast, the brightness gradient and the number of brightness levels, together with combinations of these indices such as, e.g., the integral index (see Refs. [7, 43, 44]), the index of gradient transfer [45] and the index of block similarity [46] prove to be useless for the quality assessment performed in the frame of the target-oriented IF.

In other words, most of the objective image-quality metrics available in the literature are not suitable for the assessment of the target-oriented images. We have at least two reasons for this situation:

(i) The image-quality metrics based on the comparison with a reference image cannot be used since the appropriate reference image cannot be selected among the input partial images. This is why only so-called 'no-reference' image-quality metrics can be used.

(ii) By their definition, most of the no-reference quality indices represent integral parameters which are calculated across the whole image. For this reason they prove to be rather weakly, randomly or indirectly relevant to the visibility of targets.

On the other hand, the notion of the local contrast introduced in Ref. [11] and our method suggested for evaluating the quality of the target-oriented fused images fit the corresponding needs.

It is important that the angular-azimuth (see Eq. (36)) and ellipticity (see Eq. (37)) algorithms, which can be subjected to the Jones-type transformations (see Eq. (32)), enable a dynamic fusion [14, 15]. Then the fusion algorithm can be parameterized. As a result, instead of a single fused image, one can generate a set of images for different parameter values, which can be combined into a video file. For example, one of the possible parameterization of Eq. (37) is as follows:

$$\sin 2\gamma = \frac{a\left(u^2 + \upsilon^2\right) + b\left(u\upsilon\right)}{b\left(u^2 + \upsilon^2\right) + a\left(u\upsilon\right)}. \tag{40}$$

A video fused using the algorithm given by Eq. (40) with the parameters varying in the ranges $a \in [-1; 1]$ and $b \in [1; -1]$ is shown in Fig. 4. The regime of dynamic fusion allows one to select the parameters that provide the best quality of the fused image much faster, if compared with visual scanning of the separate single images that correspond to different parameter values.

4. Conclusion

Most of the present classifications of the IF techniques have been based on the distinctive features of the image-processing procedures involved in these technique. However, the distinction among the IF techniques that follows from the aim of the IF is of no less importance. By this aim, the IF techniques can be divided into the techniques designed for art photography and those serving the activities embraced by the term 'visual search for a target'. The main aim of the target-oriented IF techniques is acquisition of some data on a target, i.e. detection, recognition and identification of the latter. In its turn, the target-oriented IF techniques can be classified into the express ones and the techniques designed for the office use.

We argue that the distinction between the express and office methods is of crucial importance for several reasons. The requirements concerned with the complexity of image-processing procedures, the processing time, the power of computers used, the computer programming qualification of users and the image-quality assessment are essentially different for these two classes of the IF methods. Our analysis shows that most of the quantitative image-quality indices available in the literature are not applicable to the task of evaluation of the target-oriented IF. In particular, the express IF methods need their specific quantitative image-quality indices selectively

associated with the target visibility and conspicuity. We suggest the so-called local contrast as a quantitative index devised for assessing the visibility of a target and present a method for its measurements.

The office IF techniques are well presented in the current literature and their number quickly increases, whereas only a few available IF techniques can be classified as the *express methods*. In the present work, we focus on the recent complex-scalar and complex-vector approaches for the express target-oriented IF. Moreover, we demonstrate that the angular-ellipticity algorithm of the CVIF technique can provide a considerably higher visibility of the target in the fused image, when compared with that obtained with the algorithms used by the other express target-oriented IF methods.

Acknowledgments

References

1. Eckstein M P, 2011. Visual search: a retrospective. J. Vis. **11**: 1–36.
2. Bowler Y M. Towards a Simplified Model of Visual Search. In: Visual Search, Ed. by D. Brogan, pp. 303–309, London: Taylor & Francis (1990).
3. Wolfe J M. Visual Search. In: The handbook of Attention. Ed. by J. M. Fawcett, E. F. Risko and A. Kingstone, pp. 27–56, Boston Review (2015).
4. Jiayi Ma, Yong Ma and Chang Li, 2019. Infrared and visible image fusion methods and applications: a survey. Inform. Fusion. **45**: 153–178.
5. Parvatikar M V and Phadke G S, 2013. Comparative study of different image fusion techniques. Int. J. Sci. Engin. Techn. **3**: 375–379.
6. Sicong Zheng. Pixel-level image fusion algorithms for multicamera imaging system. Master Thesis. University of Tennessee, Knoxville (2010).
7. Hryvachevskyi A P. Improving the informativeness of multispectral monitoring systems by image fusion of the visible and infrared ranges. Ph. D. Thesis. Lviv (2019).
8. Piella G, 2003. A general framework for multiresolution image fusion: from pixels to regions. Inform. Fus. **4**: 259–280.
9. Kaur H, Koundal D and Kadyan V, 2021. Image fusion techniques: a survey. Arch. Comp. Meth. Engin. **28**: 4425–4447.
10. Williams M L, Wilson R C and Hancock E R, 1999. Deterministic search for relational graph matching, Patt. Recog. **32**: 1255–1516.
11. Khaustov Ya Ye, Khaustov D Ye, Lychkovskyy E, Ryzhov Ye and Nastishin Yu A, 2019. Image fusion for a target sightseeing system of armored vehicles. Milit. Techn. Collect. **21**: 28–37.
12. Khaustov Ya Ye, Khaustov D Ye, Ryzhov Ye, Lychkovskyy E and Nastishin Yu A, 2020. Fusion of visible and infrared images via complex function. Milit. Techn. Collect. **22**: 20–31.
13. Khaustov Ya Ye, Khaustov D Ye, Hryvachevskyi A P, Ryzhov Ye, Lychkovskyy E, Prudyus I N and Nastishin Yu A, 2021. Complex function as a template for image fusion. Res. Opt. **2**: 100038.
14. Khaustov D Ye, Khaustov Ya Ye, Ryzhov Ye, Lychkowskyy E, Vlokh R and Nastishin Yu A, 2021. Jones formalism for image fusion. Ukr. J. Phys. Opt. **22**: 165–180.

15. Khaustov D, Khaustov Ya, Ryzhov Ye, Burashnikov O, Lychkovskyy E and Nastishin Yu, 2021. Dynamic fusion of images from the visible and infrared channels of sightseeing system by complex matrix formalism. Milit. Techn. Collect. **25**: 29–37.

16. Visual and Infrared Database Image Fusion. Provided for free downloading by the SMT/COPPE/Poli/UFRJ and IME–Instituto Militar de Engenharia within the CAPES/Pró–Defesa Program, in a partnership with IPqM–Instituto de Pesquisa da Marinha. url: http://www02.smt.ufrj.br/~fusion/

17. Li S, Kang X and Hu J, 2013. Image fusion with guided filtering. IEEE Trans. Image Process. **22**: 2864–2875.

18. Yang B and Li S, 2014. Visual attention guided image fusion with sparse representation. Optik. **125**: 4881–4888.

19. Liu Y, Liu S and Wang Z, 2015. A general framework for image fusion based on multi-scale transform and sparse representation. Inform. Fusion. **24**: 147–164.

20. Bavirisetti D P, Xiao G and Liu G, 2017. Multi-sensor image fusion based on fourth order partial differential equations, Proc. of 20th International Conference on Information Fusion. Xian, China, pp. 1–9.

21. Bavirisetti D P and Dhuli R, 2016. Two-scale image fusion of visible and infrared images using saliency detection. Infrared Phys. Technol. **76**: 52–64.

22. Xiao G, Bavirisetti D P, Liu G and Zhang X. Image Fusion based on machine learning and deep learning. In: Image Fusion. Singapore: Springer (2020).

23. 23.Yu Liu, Xun Chen, Zengfu Wang, Z Jane Wang, Raba K Ward, and Xuesong Wang, 2018. Deep learning for pixel-level image fusion: recent advances and future prospects. Inform. Fusion. **42**: 158–173.

24. Jingchun Piao, Yunfan Chen and Hyunchul Shin, 2019. A new deep learning based multi-spectral image fusion method. Entropy. **21**: 570.

25. Yu Zhang, Yu Liu, Peng Sun, Han Yan, Xiaolin Zhao and Li Zhang, 2020. IFCNN: a general image fusion framework based on convolutional neural network. Inform. Fusion. **54**: 99–118.

26. Kong W, Zhang L and Lei Y, 2014. Novel fusion method for visible light and infrared images based on nsst–sf–pcnn. Infrared Phys. Technol. **65**: 103–112.

27. Sreeja G and Saraniya O. 2019. Chap. 3. Image fusion through deep convolutional neural network in deep learning and parallel computing environment for bioengineering systems. In: Deep Learning and Parallel Computing Environment for Bioengineering Systems, Ed. by Arun Kumar Sangaiah, Elsevier, pp. 37–52.

28. Thung K-H and Raveendran P, 2009. A survey of image quality measures. 2009 International Conference for Technical Postgraduates (TECHPOS), Kuala Lumpur, Malaysia, pp. 1–4.

29. Wertheim A H, 2010. Visual conspicuity: a new simple standard, its reliability, validity and applicability. Ergonomics. **53**: 421–442.

30. Huckauf A, Heller D and Nazir T A, 1999. Lateral masking: limitations of the feature interaction account. Percept. Psychophys. **61**: 177–189.

31. Krikke K, Wertheim A H and Johnson A, 2000. The role of lateral masking in visual search. Perception (Suppl.). **29**: 102.

32. Wertheim A H, Hooge I T C, Krikke K and Johnson A, 2006. How important is lateral masking in visual search? Exp. Brain Res. **170**: 387–402.

33. Wolford G and Chambers L, 1983. Lateral masking as a function of spacing. Percept. Psychophys. **33**: 129–138.

34. Westheimer G, 1985. The oscilloscopic view: retinal illuminance and contrast of point and line targets. Vision Res. **25**: 1097–1103.

35. Adrian W, 1989. Visibility of targets: model for calculation. Light. Res. Technol. **21**: 181–188.

36. Landsberg G S. Optics. Moscow: Nauka (1976).

37. Born M and Wolf E. Principles of Optics. Electromagnetic Theory of Propagation, Interference and Diffraction of Light. 7th Ed. Cambridge: Cambridge University Press (2005).

38. Azzam R M A and Bashara N M, Ellipsometry and Polarized Light. Amsterdam, New York: North-Holland Publ. (1977).

39. Nastyshyn S Yu, Bolesta I M, Tsybulia S A, Lychkovskyy E, Yakovlev M Yu, Ryzhov Ye, Vankevych P I and Nastishin Yu A, 2018. Differential and integral Jones matrices for a cholesteric. Phys. Rev. A. **97**: 053804.

40. Nastyshyn S Yu, Bolesta I M, Tsybulia S A, Lychkovskyy E, Fedorovych Z Ya, Khaustov D Ye, Ryzhov Ye, Vankevych P I and Nastishin Yu A, 2019. Optical spatial dispersion in terms of Jones calculus. Phys. Rev. A. **100**: 013806.

41. Konstantinova A F, Grechushnikov B N, Bokut B V and Valyashko Ye G. Optical properties of crystals. Minsk: Navuka i Teknnika (1995).

42. J. Eli Peli, 1990. Contrast in complex images. Opt. Soc. Amer. A. **7**: 2032–2040.

43. Bogdanov P and Romanov Yu N, 2012. Quality assessment of digital images. Mechanics, Control and Informatics. **9**: 218–226.

44. Bondarenko M A, Drynkin V N, Nabokov C A and Pavlov Yu V, 2017. Adaptive algorithm for selecting informative channels in onboard multispectral video systems. Prog. Syst. Comp. Meth. **1**: 46–52.

45. Xydeas C S and Petrović V, 2000. Objective Image Fusion Perf. Meas. Electron. Lett. **36**: 308–309.

46. Cvejic N A, Loza D and Bul N, 2005. Similarity metric for assessment of image fusion algorithms. Int. J. Sign. Proc. **2**: 178–182.

Анотація. *Орієнтовані на пошук цілей (або «цільові») експрес-методи злиття зображень (ЗЗ) спрямовані на отримання даних про ціль, тобто на її візуальний пошук, виявлення, розпізнавання та ідентифікацію. Вони формують окремий клас методів ЗЗ, відмінний від класу методів, орієнтованих на художню фотографію. Найбільш очевидним аргументом на користь їхнього розрізнення є те, що вимоги до складності процедур обробки зображень та оцінки їхньої якості істотно відрізняються для цих класів. Зокрема, ці вимоги мають першочергове значення для одного із вищезазначених класів і не мають значення (або навіть не застосовні) для іншого класу. Цільові експрес-методи ЗЗ потребують власних кількісних показників якості зображення, які повинні бути вибірково пов'язаними саме з видимістю чи помітністю цілі. Після обговорення особливостей оцінки якості зображень, злитих за експрес-методами цільового ЗЗ, запропоновано спосіб визначення локального контрасту цілі та встановлено зв'язок цього показника з видимістю цієї цілі. Ми зосередилися на застосуванні комплексно-скалярного та комплексно-векторного підходів ЗЗ, які були розроблені останнім часом, до вирішення проблем експрес-методів цільового ЗЗ. Показано, що алгоритм еліптичності комплексно-векторного ЗЗ забезпечує значне підвищення видимості цілі, порівняно з іншими експрес-методами цільового ЗЗ.*

Ключові слова*: злиття зображень, збір даних про ціль, метод комплексно-скалярного злиття зображень, метод комплексно-векторного злиття зображень, локальний контраст цілі, видимість цілі*