



Reconnaissance hors-contact du stress mental en utilisant la variabilité cardiaque

Frédéric Bousefsaf, Choubeila Maaoui, Alain Pruski

► To cite this version:

Frédéric Bousefsaf, Choubeila Maaoui, Alain Pruski. Reconnaissance hors-contact du stress mental en utilisant la variabilité cardiaque. Colloque Jeunes Chercheurs, Jeunes Chercheuses IFRATH, 2013, Paris, France. hal-03790882

HAL Id: hal-03790882

<https://hal.science/hal-03790882>

Submitted on 28 Sep 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Reconnaissance hors-contact du stress mental en utilisant la variabilité cardiaque

Frédéric Bousefsaf, Choubeila Maaoui, Alain Pruski

Laboratoire de Conception, Optimisation et Modélisation des Systèmes (LCOMS)
7 rue Marconi, 57070 Metz, France
{frederic.bousefsaf, choubeila.maaoui, alain.pruski}@univ-lorraine.fr

Résumé. Dans ce papier, nous présentons un ensemble de méthodes qui permettent de détecter des états de stress mental en utilisant des trames vidéo obtenues à partir d'une webcam bas coût. Les mesures sont réalisées spécifiquement sur les visages des participants et les algorithmes que nous proposons ont été développés de manière à ce que l'application soit robuste aux bruits générés par les mouvements de la tête. Une version interactive du test de Stroop a été employée pour induire du stress chez les 12 candidats qui ont participé à cette étude. Parallèlement, nous avons enregistré la conductance de la peau que nous avons comparée à la courbe de stress obtenue par une analyse sur la variabilité cardiaque de la webcam. Les résultats apportent une nouvelle manière de détecter des états de stress en utilisant des moyens bas coûts, offrant une alternative aux techniques conventionnelles en contact.

Mots-clés : variabilité cardiaque; hors contact; émotions; stress mental; réalité virtuelle; webcam

1 Introduction

Reconnaître l'état émotionnel d'un patient est particulièrement utile dans le cadre des thérapies basées sur la Réalité Virtuelle (RV) où les émotions sont la boucle de retour qui permet de réguler l'intensité de l'environnement virtuel [1]. Ces scènes virtuelles basées sur la physiologie sont particulièrement utilisées pour traiter des personnes atteintes de troubles anxieux et de phobies, comme les soldats ou les anciens combattants [2]. Certains signaux physiologiques, comme le rythme cardiaque (RC) et la Variabilité de la Fréquence Cardiaque (VFC) sont reconnus pour être des paramètres fiables dans le domaine de la reconnaissance des émotions [3]. Cependant, les capteurs en contact peuvent présenter des limites dans certains champs applicatifs où un spécialiste doit être physiquement présent pour poser et contrôler le bon fonctionnement des capteurs. Ces derniers peuvent gêner l'utilisateur lors du déroulement d'une activité, comme par exemple lors de situations de réhabilitation motrice ou encore pendant l'utilisation de serious games. Les capteurs en contact peuvent aussi être un frein quand ils deviennent intrusifs pour l'intimité du patient.

La mesure hors contact des signaux physiologiques peut être réalisée en utilisant des caméras thermiques infrarouges. Employée par Pavlidis et al., cette technologie permet de mesurer des données physiologique en se focalisant sur le visage [4]. De la même manière, les radars basés sur l'effet Doppler sont des capteurs hors contact qui permettent de détecter les battements du cœur [5] et la respiration [6] d'un individu. Plus récemment, les caméras numériques et les webcams ont été utilisées pour mesurer l'activité cardiaque [7] sur des visages humains afin d'en extraire les rythmes cardiaque et respiratoire. Le principe, basé sur la PhotoPlethysmoGraphie (PPG) consiste à observer les variations de lumière sur la peau du visage pour en extraire le flux sanguin et ainsi délivrer une image de l'activité cardiaque. Sun et al. [8] ont comparé les performances en termes d'extraction du RC entre une caméra numérique haute sensibilité et une webcam bas coût. Leur conclusion indique que les caractéristiques de la webcam échantillonnée à 30 fps sont comparables à celles de la caméra échantillonnée à 200 fps après avoir interpolé les signaux pour améliorer la résolution temporelle [8].

La VFC est un paramètre de santé important, utilisé également en informatique affective et en psychophysiologie pour donner une image de l'activité du Système Nerveux Autonome (SNA) afin de détecter des états de stress mental en temps réel [9]. L'analyse spectrale du SNA nous informe sur la balance sympathovagale, un index qui reflète les variations mutuelles des systèmes nerveux sympathique et parasympathique [10]. La VFC est régulière et rythmée lorsque l'individu est au calme et suit la respiration par un phénomène appelée Arythmie Sinusale Respiratoire. Au contraire, sa forme et son allure deviennent chaotiques lorsque la personne endure un stress ou une anxiété intense. L'état, donné par ces variations du rythme cardiaque, est appelée cohérence cardiaque [11]. L'extraction des signaux physiologiques par des technologies hors-contact est particulièrement intéressante pour des applications qui cherchent à comprendre et synthétiser les émotions et sentiments d'un patient. L'état émotionnel peut faire office de boucle de retour et est employé dans des applications basées sur la réalité virtuelle, comme les thérapies d'exposition par RV [12]. Le control en temps réel des signaux physiologiques permet au thérapeute qui supervise la session d'exposition de paramétrer et d'optimiser les stimuli, situations et environnements virtuels. Durant ces séries de thérapies psychophysiologiques, les capteurs en contact peuvent générer un biais en gênant le patient, ce qui se caractérise en pratique par une estimation erronée des émotions [4].

Des algorithmes basés sur l'apprentissage automatique peuvent être utilisés sur les paramètres extraits des signaux physiologiques pour automatiquement estimer l'état émotionnel [3], [13] et spécifiquement pour les états de stress [14]. Des études ont montré que seulement la VFC est utilisable pour détecter des états de stress mental [9], [15], [16]. L'ajout de modalités [17] permet d'augmenter la qualité et le nombre d'émotions reconnues en comparaison avec les approches qui reposent sur une unique modalité.

Dans cette étude, nous proposons une méthode originale qui permet d'extraire le RC instantané à partir de trames vidéo contenant des visages humains en utilisant une webcam HD bas coût. La VFC est calculée et utilisée pour reconnaître l'état de stress mental des 12 participants qui ont participé à cette étude, pendant qu'ils utilisaient une version interactive du test de Stroop.

2 Méthodes

2.1 Extraction du flux sanguin à partir des trames vidéo

La méthode présentée dans cette étude est composée d'opérations de traitement du signal et des images (Fig. 1). Nous utilisons une caméra qui délivre des images HD, augmentant ainsi le nombre de pixels qui composent le visage et qui contiennent l'information PPG. L'acquisition et le traitement d'images HD vont parallèlement augmenter les temps de calcul. Nous avons utilisé les paramètres Pan, Tilt et Zoom de la webcam afin de garantir un temps de rafraîchissement de 30 trames par seconde. Le visage est automatiquement détecté en utilisant une cascade de classifieurs boostés sur chaque frame via la librairie OpenCV 2.1. Les paramètres Pan et Tilt sont calculés pour suivre la position du visage dans le temps, tandis que le Zoom est paramétré pour que le visage remplisse au maximum la trame d'entrée [Fig. 1 (a)]. Des opérations de prétraitement sont appliquées pour isoler les pixels de la couleur de la peau qui contiennent l'information PPG. Un filtre qui détecte la peau [Fig. 1 (b)] est utilisé pour collecter ces pixels.

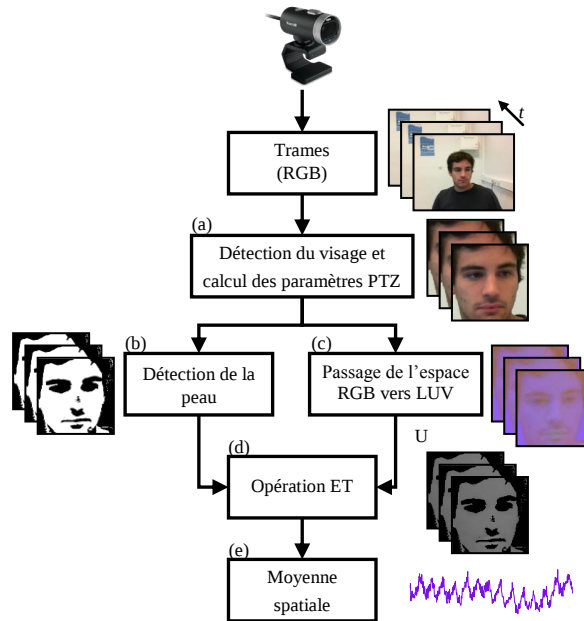


Fig. 1. Vue d'ensemble de la méthode de détection du signal PPG. (a) Les paramètres Pan, Tilt et Zoom sont calculés pour suivre le visage. (b) Les pixels qui contiennent l'information PPG sont isolés par un filtre qui détecte la peau. (c) L'espace RGB est converti vers l'espace CIE LUV. (d) La trame U de l'espace LUV est combinée à la trame de la détection de la peau. (e) Une moyenne spatiale est employée pour transformer une série de trames en un signal brut.

Des opérations de prétraitement sont appliquées pour isoler les pixels de la couleur de la peau qui contiennent l'information PPG. Un filtre qui détecte la peau [Fig. 1 (b)] est utilisé pour collecter ces pixels. L'espace colorimétrique CIE LUV a été choisi dans cette étude [Fig. 1 (c)] pour atténuer les fluctuations dues aux artefacts de mouvement et de lumière. L'oxyhémoglobine et la désoxyhémoglobine ont un meilleur coefficient d'absorption par la peau pour des longueurs d'onde comprises entre 540 et 577 nm, ce qui correspond globalement à la couleur jaune/verte [18]. L'information PPG est principalement influente sur la composante U de l'espace colorimétrique CIE LUV qui représente les couleurs comprises entre le rouge et le vert. Une moyenne spatiale est calculée sur les intensités des pixels de la composante U qui existent sur la trame de détection de la peau. Un ensemble de n trames donne un signal brut de n point [Fig. 1 (e)]. Ce signal est traité en utilisant un filtrage sur la transformée en ondelettes continue, spécifiquement pour supprimer les tendances et bruits haute fréquence [Fig. 2 (a)] dans la bande fréquentielle 0.65–3 Hz. Un algorithme a été développé pour détecter les pics et calculer le RC instantané du participant [Fig. 2 (c)].

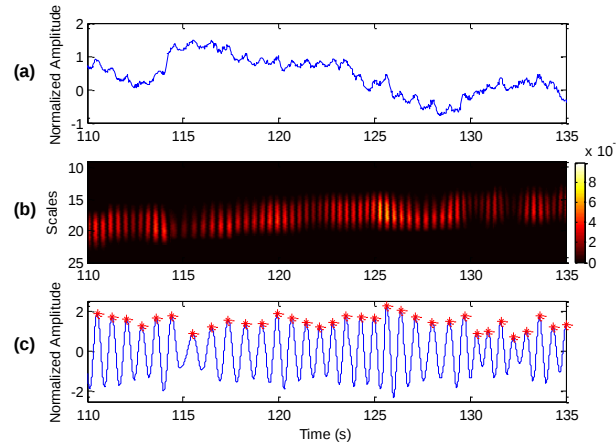


Fig. 2. Traitement du signal. Le signal brut (a) est filtré en utilisant sa transformée en ondelettes continue (b). Ce tracé est un scalogramme représentant le pourcentage d'énergie pour chaque coefficient. Les tendances et bruits haute fréquence du signal filtré (c) sont supprimées en calculant la transformée inverse dans la bande fréquentielle 0.65–3 Hz.

2.2 Une application à la détection de l'effort mental

2.2.1 Procédure

Douze étudiants (deux femmes et dix hommes) du laboratoire ont participé à cette étude. Tous les participants nous ont donné leur consentement avant le démarrage de l'expérience. Cette dernière dure de cinq à six minutes. La tâche informatique que nous proposons a d'ores et déjà été utilisée dans d'autres études.

Cette dernière est basée sur une version interactive du test colorimétrique de Stroop [15]. Très brièvement, le participant a 3 secondes pour cliquer sur la zone colorée rectangulaire qui correspond à la couleur indiquée par le mot imprimé au centre de l'écran [Fig. 4 (a)]. Certains mots sont affichés dans une couleur différente du nom qu'elle indique (incongrus, comme par exemple le mot "vert" affiché en couleur bleu) tandis que les autres sont affichés dans la couleur correspondante (congrus, comme par exemple le mot "rose" affiché en couleur rose).

Les participants effectuent trois sessions (voir Fig. 3) du test interactif de Stroop. Ils démarrent avec une session d'apprentissage (SA) d'une minute afin de leur permettre de s'habituer à l'interface et au jeu virtuel. Deux sessions de stress (SS) d'une minute chacune sont lancées par la suite. Chaque SS est séparée par une session de relaxation (SR) correspondant à une vidéo d'une minute basée sur la nature sauvage où les animaux. Dans la première SS, le participant a une minute pour cliquer correctement sur 35 zones rectangulaires. Une erreur lance automatiquement un son d'erreur et décrémente le compteur d'une unité. Une barre de progression horizontale est ajoutée en dessous du mot affiché au milieu de l'écran pour renseigner l'utilisateur sur le temps restant avant la fin de la session. De plus, une barre de progression verticale affichée à droite du mot indique le temps restant pour cliquer sur la zone colorée correspondante. La seconde SS est identique à la première, mis à part pour les positions des zones de couleurs qui se mélangent aléatoirement à chaque click juste. Cette fois, le participant doit cliquer correctement sur 40 zones. Une musique stressante est jouée pendant les deux SS et une sirène d'alarme est lancée les 10 dernières secondes.

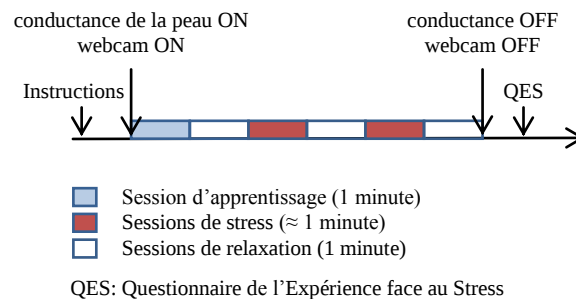


Fig. 3. Protocol expérimental.

A la fin de l'expérience, nous avons demandé aux participants de remplir un questionnaire lié à leur expérience par rapport au stress généré par l'application en utilisant une échelle de Likert sur 5 points (1 = pas du tout, 5 = Tout à fait). Les cinq paramètres suivants apparaissent dans le questionnaire : étiez-vous stressé, tendu, épuisé, concentré et stimulé. Les participants répondent ainsi deux fois à cette batterie de questions : une fois pour les sessions de stress [Fig. 4 (a)] et une fois pour les sessions de relaxation [Fig. 4 (b)]. Cette notation est utilisée pour contrôler la corrélation entre les réponses physiologiques et leur perception de l'effort mental.

Finalement, une dernière question leur est posée pour examiner les effets du mélange aléatoire des positions des zones pendant la deuxième SS. Les réponses électrodermales (RED) des participants ont été simultanément enregistrées en mesurant la conductance de la peau à l'aide d'un capteur en contact (SC-Flex/Pro by Thought Technologies Ltd.).



Fig. 4. Captures d'écran de l'application : La première session de stress présentant un cas incongru (a) et la première vidéo de relaxation (b) qui démarre directement après la session d'apprentissage.

2.2.2 Analyse de la VFC pour détecter des états de stress

Les opérations de traitement du signal détaillées dans la section 2.1 ont été utilisées pour extraire le RC instantané [Fig. 5 (a)] des participants. Les signaux bruts irrégulièrement distribués sur l'axe des temps sont ré-échantillonnés par une spline cubique sur 15 Hz. Le RC moyen augmente lors des périodes stressantes. Pour détecter ces variations de tendances, une moyenne glissante sur 400 points (27 secondes) est calculée à partir du signal interpolé du RC instantané [Fig. 5 (b)]. L'allure de la VFC est rythmée et régulière pendant les SR et chaotique et irrégulière pendant les SS [11]. Pour détecter ces fluctuations, nous utilisons la dérivée troisième du signal interpolé [Fig. 5 (c)]. De plus, les variations en fréquence sont détectées [Fig. 5 (d)] par un algorithme basé sur l'analyse haute fréquence de la transformée en ondelettes continue du signal interpolé.

Ces trois paramètres sont ensuite additionnés et le résultat est lissé par une moyenne glissante sur 400 points, délivrant ainsi une unique représentation de l'état de stress par une courbe (voir Fig. 6). Les points maximums et minimums sont détectés, ce qui correspond en pratique à des maximums de stress et de relaxation respectivement.

3 Résultats

Les courbes de stress ont été comparées aux RED afin de mesurer la corrélation avec les mesures de la caméra. Un exemple est présenté sur la Fig. 6 où la courbe de stress généré par la VFC est en étroite corrélation avec la RED.

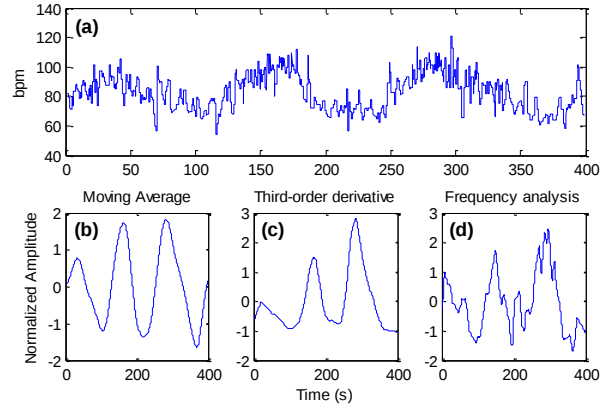


Fig. 5. Résultats de l'analyse sur la VFC d'un participant (a). Une moyenne glissante (b) est calculée sur la VFC interpolée pour détecter les variations de tendances. La dérivée troisième (c) est utilisée pour détecter les fluctuations de rythme sur la variabilité. Une analyse sur les hautes fréquences de la transformée en ondelettes continue de la VFC (d) donne une image du SNA.

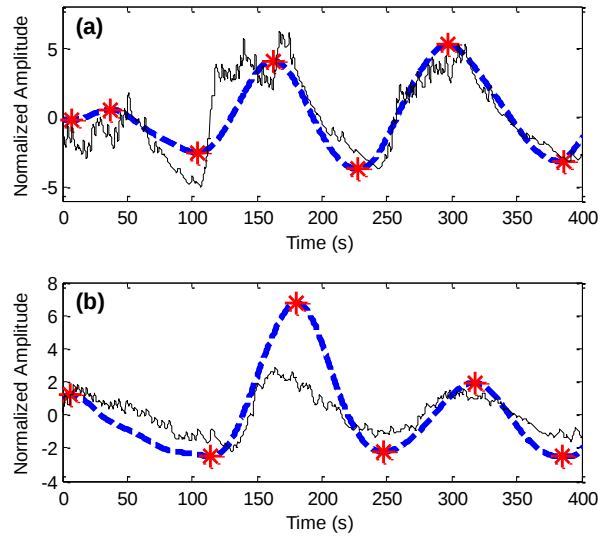


Fig. 6. Résultats de la détection du stress pour le participant #1 (a) et #8 (b). La courbe tracée avec des tirets correspond au signal de stress dérivé de la webcam tandis que celle en trait plein correspond à la RED du participant.

Des diagrammes de Tukey (boîtes à moustache) sont employés pour représenter les différences au niveau de la moyenne de la courbe de stress et des RED pour chaque session (Fig. 7). Des différences significatives sont identifiables sur les réponses des questionnaires (Tableau 1) entre les SR et les SS en suivant les facteurs *stressé*, *tendu*, *concentré* et *stimulé*. A l'inverse et toujours selon les réponses du questionnaire, le facteur *épuisé* ne présente pas de variation significative entre les différentes sessions.

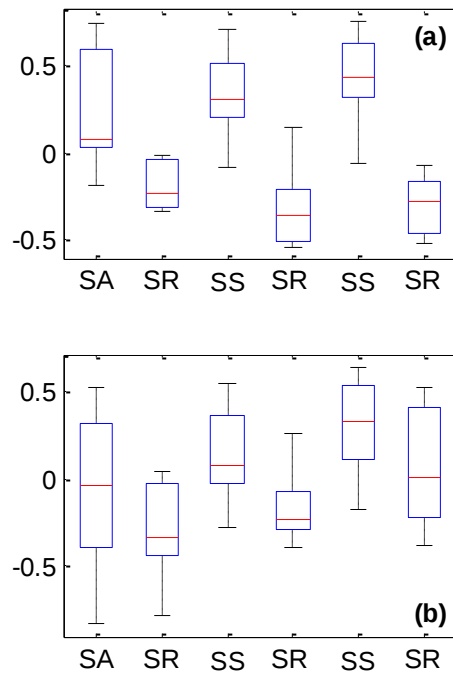


Fig. 7. Diagrammes de Tukey représentant les valeurs moyennes des courbes pour chaque session. Les mesures de la webcam (a) et les RED (b) indiquent que les moyennes augmentent significativement pendant les deux sessions de stress comparées aux moyennes des sessions de relaxation. Ces boîtes englobent les données des 12 participants de l'étude.

4 Discussion et conclusion

Détecter le RC et la VFC offre une possibilité hors contact pour mesurer l'état de stress d'un individu tout en contrôlant d'importants paramètres physiologiques. Nous avons choisi de travailler avec des webcams, une technologie accessible, pour mesurer ces données. Les résultats présentés dans cette étude démontrent qu'il est possible d'utiliser les réponses de la VFC à partir d'une webcam bas coût pour détecter des états d'effort mental. Les opérations de traitement du signal et des images présentées dans

cette étude sont robustes aux conditions d'illumination de la pièce et tolérantes face au mouvement. Les diagrammes de Tukey illustrés en Fig. 7 donnent une estimation de l'état de la courbe de stress des participants. Ces courbes tendent à décroître, identiquement à l'allure des RED, pendant les phases de relaxations. Au contraire, les tendances de ces courbes vont progressivement augmenter pendant les sessions de stress. La disparité sur les diagrammes de la webcam [Fig. 7 (a)] est relativement élevée et nous indique que les participants étaient en moyenne plus stressés pendant la SA et les SS que pendant les SR. Ces données sont en concordance avec les réponses aux questionnaires qui reflètent l'expérience subjective des participants par rapport au stress.

Tableau 1. Résultats des questionnaires relatifs au stress.

Variable	SR	SS
Stressé	1.92 (0.99)	3.75 (0.75)
Tendu	1.58 (0.90)	3.58 (0.90)
Epuisé	1.25 (0.62)	1.83 (0.83)
Concentré	2.75 (1.42)	4.58 (0.67)
Stimulé	2.92 (1.38)	3.92 (0.79)

Ces valeurs représentent la moyenne (écart type) pour chaque facteur sur chaque participant, de 1 (pas du tout) à 5 (tout à fait).

Les clicks erronés ont été enregistrés pendant les expériences et devront être analysés dans des travaux futurs pour observer leurs impacts sur l'effort et le stress mental des participants. Le nombre d'émotions reconnues peut être augmenté en intégrant plus de modalités à l'application. Les vidéos peuvent ainsi être analysées pour proposer un système de reconnaissance des émotions multimodal. La capacité de comprendre et de détecter les émotions humaines pour les thérapies basées sur la RV permet de réguler le niveau et l'intensité des environnements virtuels proposés par les thérapeutes.

5 Bibliographie

1. Kukulja D., Popovic S., Dropuljic B., et al.: Real-time emotional state estimator for adaptive virtual reality stimulation. In: Proc. 5th Int. Conf. Foundations of Augmented Cognition. Neuroergonomics and Operational Neuroscience, pp. 175--184. San Diego, CA, USA (2009)
2. Rizzo A.A., Difede J., Rothbaum B.O., et al.: VR PTSD exposure therapy results with active duty OIF/OEF combatants. *Medicine meets Virtual Reality* 17, 277--282 (2009)
3. Jerritta S., Murugappan M., Nagarajan N, and Wan K.: Physiological signals based human emotion Recognition: a review. In: IEEE 7th Int. Colloq. Signal Process. Appl., pp. 410--415. Arau, Malaysia (2011)
4. Pavlidis I., Dowdall J., Sun N., et al.: Interacting with human physiology. *Computer Vision and Image Understanding* 108, 150--170 (2007)

5. Vasu V., Heneghan C., Arumugam T., and Sezer S.: Signal processing methods for non-contact cardiac detection using Doppler radar. In: IEEE Workshop on Signal Process. Syst., pp. 368--373. San Francisco, CA (2010)
6. Gu C., Li R., Li C., and Jiang S.B.: Doppler radar respiration measurement for gated lung cancer radiotherapy. In: IEEE Topical Conf. Biomed. Wireless Technol. Netw. Sens. Syst., pp. 91--94. Phoenix, AZ (2011)
7. Poh M.-Z., McDuff D.J., and Picard R.W.: Advancements in noncontact, multiparameter physiological measurements using a webcam. IEEE Trans. Biomed. Eng. 58, 7--11 (2011)
8. Sun Y., Hu S., Azorin-Peris V., et al.: Noncontact imaging photoplethysmography to effectively access pulse rate variability. J. Biomed. Opt. 18, 061205 (2013)
9. Hoover A., Singh A., Fishel-Brown S., and Muth E.: Real-time detection of workload changes using heart rate variability. Biomedical Signal Processing and Control 7, 333--341 (2012)
10. Appelhans B.M. and Luecken L.J.: Heart rate variability as an index of regulated emotional responding. Review of General Psychology 10, 229--240 (2006)
11. Tiller W.A., McCraty R. and Atkinson M.: Cardiac coherence: a new, noninvasive measure of autonomic nervous system order. Altern. Ther. Health Med. 2, 52--65 (1996)
12. Cosic K., Popovic S., Kukolja D., et al.: Physiology-driven adaptive virtual reality stimulation for prevention and treatment of stress related disorders. Cyberpsychol. Behav. Soc. Netw. 13, 73--78 (2010)
13. Knapp R.B., Kim J., and André E.: Physiological signals and their use in augmenting emotion recognition for Human-Machine Interaction. In: Emotion-Oriented Systems, Springer-Verlag, pp. 133--159 (2011)
14. Zhai J. and Barreto A.: Stress recognition using non-invasive technology. In: Proc. 19th Int. Florida AI Res. Soc. Conf., pp. 395--401, Melbourne Beach, Florida, USA (2006)
15. Garde A.H., Laursen B., Jorgensen A.H., and Jensen B.R.: Effects of mental and physical demands on heart rate variability during computer work. Eur. J. Appl. Physiol. 87, 456--461 (2002)
16. Hercegfi K.: Heart rate variability monitoring during Human-Computer Interaction. Acta Polytechnica Hungarica 8, 205--224 (2011)
17. Gonzalez-Sanchez J., Chavez-Echeagaray M.E., Atkinson R., Burleson W.: ABE: an agent-based software architecture for a multimodal emotion recognition framework. In: Proc. 9th Working IEEE/IFIP Conf. Softw. Archit., Boulder, Colorado, USA (2011)
18. Horecker B.L.: The absorption spectra of hemoglobin and its derivatives in visible and near infrared regions. J. Biol. Chem. 148, 173--183 (1943)