

Utilisation de techniques de réalité mixte dédiées au docking peptide/protéine.

Jeux sérieux pour l'enseignement et la recherche

Directeur de thèse : Patrick Bourdot

Co-direction : Jean-Marc Vézien

Using mixed-reality techniques dedicated to peptid/protein docking

Serious game for teaching and research

Keywords : *Markerless Tracking, Serious Game, Tangible Interface, Molecular Simulation, High Performance Computing*

Résumé :

Les interactions protéine-peptide sont essentielles pour la cellule. Les peptides ont un rôle longtemps sous-estimé, pourtant, ils négocient, inhibent ou servent d'éléments structurant dans près de 40% de toutes les interactions macromoléculaires. Ils sont souvent associés à des maladies, et constituent aussi une nouvelle piste en ingénierie protéique pour la conception de médicaments. Nous proposons d'adresser cette problématique par une approche de jeu sérieux, dont la finalité n'est pas le grand public, mais la recherche et l'enseignement. Dans ce jeu sérieux, l'étudiant ou le scientifique devra construire et manipuler une interface tangible déformable du peptide, qui sera trackée pour capturer durant la construction et manipulation, la configuration de l'objet tangible déformable. Ce tracking temps réel permettra de reconstruire un modèle virtuel de l'objet manipulé dans son environnement protéique. Il sera visualisé conjointement avec le modèle réel, complétant ainsi la représentation atomique de l'interface tangible, par d'autres représentations moléculaires classiques. Par ailleurs, pour évaluer la qualité de la solution durant la construction de son modèle physique de peptide, différentes énergies seront présentées visuellement à l'utilisateur comme scores biologiquement pertinents, calculées en temps réels par des outils existants de simulation moléculaire. Selon le contexte d'enseignement ou de recherche, le but du jeu sérieux sera respectivement d'atteindre un score maximal, correspondant à une solution de repliement ou de docking optimale proposée par le joueur, ou pour explorer de manière interactive le paysage énergétique des différentes conformations construites par le joueur.

Abstract :

Protein-peptide interactions are essential for the living cell. For a long time, the role of peptides has been underestimated, however they negotiate, inhibit or act as structuring elements in nearly 40% of all macromolecular interactions. They are often associated with diseases, and also provide a new track in protein engineering and more generally in drug design. We propose to address protein-peptide interactions by following an approach based on serious games, targeting the research and teaching communities. In this serious game, the student or scientist will be able to build and manipulate a deformable tangible interface of the peptide, which will be tracked in 3D to capture the configuration of the deformable tangible object during construction and manipulation. This real-time tracking will thus aim to reconstruct a virtual model of the real manipulated object in its proteic environment, so as to complete the atomic representation of the tangible interface with computer-generated molecular representations. Furthermore, in order to assess the quality of the solution during the construction of the physical peptide model, energies involved in the molecular system will be computed in real time using existing molecular simulation tools and displayed as biologically relevant scores. Depending on the context of teaching or research, the goal of such a serious game will be respectively to reach a maximum score corresponding to an optimal refolding or docking solution proposed by the player, or to explore the configuration landscape of the molecular system in order to find optimized solutions using the user expertise.

Le contexte scientifique

Les interactions protéine-peptide sont essentielles pour la cellule. Les peptides ont un rôle longtemps sous-estimé, pourtant, ils négocient, inhibent ou servent d'éléments structurants dans près de 40% de toutes les interactions macromoléculaires. Ils sont souvent associés à des maladies, et constituent aussi une nouvelle piste en ingénierie protéique pour la conception de médicaments (*Drug Design*). Au cours des dernières années, des technologies à grande échelle ont permis des études approfondies sur les préférences de reconnaissance de peptide pour un certain nombre de familles de domaines peptidiques de liaison [12]. Il reste cependant que le manque de données concernant les mécanismes d'interactions moléculaires avec leur partenaire, du fait de leur flexibilité, rend la prédiction de structure protéine-peptide très difficile et complexe. Cela nécessite de repenser les approches de modélisation classique utilisée en docking protéine/protéine, pour prendre en compte la flexibilité intrinsèque de ces peptides, tout en intégrant mieux l'expertise humaine dans la construction de ces modèles théoriques, pour de mieux prendre en compte les résultats expériences biologiques, et plus largement les connaissances de l'expert.

Les approches de docking protéine/protéine automatique classique comme [11], nécessitent l'implication de l'expertise humaine, exclusivement en début et en fin de chaîne. En début de chaîne, en imposant des contraintes en rapport avec les résultats biologiques ou la connaissance que l'expert a du système étudié, et en fin de chaîne, en sélectionnant parmi les nombreuses solutions proposées par le programme la plus pertinente biologiquement. C'est la raison pour laquelle des travaux concevant et évaluant des approches de docking plus interactif ont été proposés [8], pour mieux intégrer l'expertise humaine, et dans une boucle plus courte. C'est ce type de nouvelles approches de visualisation et d'interaction 3D pour l'étude et la compréhension des structures moléculaires complexes, que la *Faraday Discussion* [1], une institution centenaire importante dans la communauté de la chimie, et supportée par la *Royal Society of Chemistry*, a souhaité mettre en valeur. Cette conférence a réunit les pionniers dans la conception des interfaces avancées dédiées à la biologie structurale, comme *Frederick P. Brooks* et *Arthur J. Olson*. Il a été souligné l'importance de l'émergence de nouvelles technologies, désormais accessibles, de visualisation et d'interaction 3D en biologie structurale, étant donné la complexité croissante et la dynamique des structures moléculaires abordées aujourd'hui, grâce à d'importants progrès expérimentaux (notamment SAXS et CryoEM).

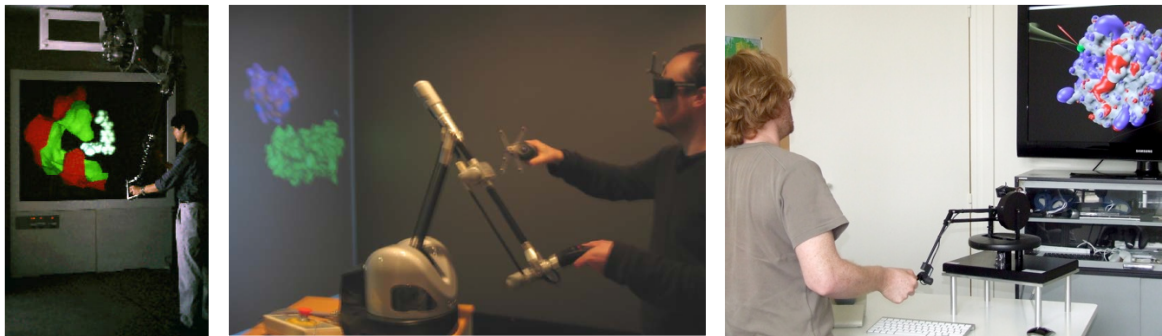


Figure 1 : A gauche : premier dispositifs haptiques dédiés au docking moléculaire (*Frederick P. Brooks*). Au milieu : utilisation de la multi-sensorialité (visuel, haptique et audio) pour le docking protéine protéine dans le projet ANR CorSAIRe (*P. Bourdot*) [8]. A droite : utilisation de technique de simulation moléculaire interactive qui consiste à appliquer interactivement des contraintes à une simulation moléculaire en cours (*M. Baaden*) [2], l'utilisateur étant lui aussi contraint par les force en jeu dans la simulation.

C'est *Frederick P. Brooks* qui a initié ce travail non seulement en informatique, mais aussi dans la manipulation interactive de structure protéique grâce à l'utilisation d'interface haptique (voir Figure 1 - à gauche), excessivement couteuses et volumineuses dans les années 70, mais aujourd'hui disponibles à bas coût et utilisables quotidiennement dans un environnement de bureau ou de réalité virtuelle. C'est dans la lignée des travaux de *Brooks* que certains travaux plus récents ont étudié les aspects multimodaux des interactions homme-machine dédiés au docking protéine/protéine comme dans [8], mais aussi appliqués avec succès pour résoudre des problématiques biologiques complexes [2]. Cependant, ces techniques de simulations moléculaires interactives montrent leurs limites, en particulier pour manipuler des modèles protéiques virtuels dynamiques et donc flexibles, essentiellement dues aux limites des 6 degrés de liberté en entrée de la plupart des dispositifs d'interaction 3D. En effet, même si il est désormais possible du fait

des progrès en simulation numérique, d'accéder par simulation à la dynamique des structures protéiques en temps interactif, il reste que les interfaces homme-machine classiques, qui parvenaient aisément à manipuler des modèles rigides, peinent à fournir des techniques d'interaction adaptées pour appréhender la dynamique des molécules *in virtuo*.

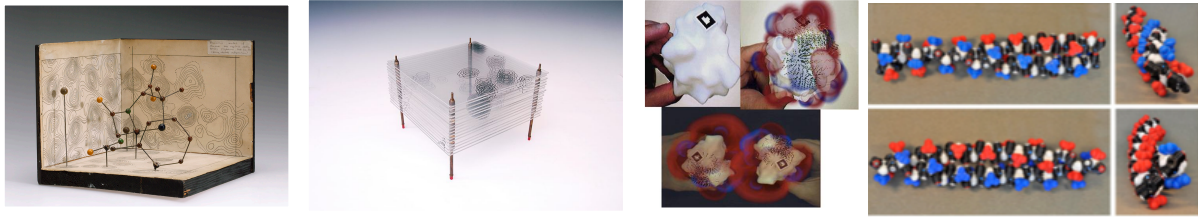


Figure 2 : De gauche à droite : modèle de la pénicilline, on peut souligner l'énorme travail lié aux plans de coupe, générés à partir du calcul (à la main) puis du dessin (à la main) de la densité électronique pour chaque point des trois plans, technique désormais très classique en visualisation scientifique (en 1945, par D. Hodgkin) - préfiguration de technique de visualisation de type *volume rendering* de la densité électronique utilisant des plans de coupes parallèles transparents sur le même modèle (en 1945, par D. Hodgkin) - interface tangible rigide d'une protéine imprimée complétée par des rendus visuels utilisant des techniques de réalité augmentée (A Gilet, A.J. Olson) [10]. Interface tangible protéique ou peptidique, modulaire et déformable, construite à base de composants identiques, et permettant de ressentir un feedback haptique à l'aide d'aimants, en rapport avec les énergies de torsion réalistes au sein de la chaîne polypeptidique [3], et représentant aussi les liaisons H. Différentes structures secondaires protéiques (feuille, boucle, hélice) peuvent être obtenues à l'aide de cette interface tangible déformable.

C'est la raison pour laquelle l'équipe d'Arthur J. Olson commença à travailler sur l'utilisation d'interfaces tangibles dédiées à la biologie moléculaire, combinée à des techniques de réalité augmentée. Il fut un précurseur dans ce domaine [9,10], son approche étant facilitée par l'émergence des techniques d'impression 3D et de nouvelles techniques de tracking plus performantes. Historiquement, les modèles physiques, interfaces tangibles avant la lettre, font en effet partie de la culture historique et scientifique de la biologie structurale et de la biophysique, car les phénomènes étudiés sont inaccessibles à la perception humaine (voir Figure 2, ci-dessus). De nos jours, les interfaces tangibles permettent de se passer du média numérique durant l'interaction, les dispositifs d'interaction étant parfois plus difficilement accessibles aux biologistes. Elles permettent aussi de bénéficier des avantages de la manipulation bi-manuelle naturelle d'objet physique, et qui, combinée avec des techniques de réalité augmentée, donne accès à des rendus visuels complémentaires à la représentation physique (voir Figure 2).

Projet de recherche

Nous proposons de développer un jeu sérieux dans lequel l'utilisateur devra construire et manipuler une interface tangible déformable du peptide, qui sera trackée pour capturer durant la construction et la manipulation, la configuration de l'objet tangible déformable. L'utilisation d'interfaces tangibles déformables, assemblées et manipulées par l'opérateur au fur et à mesure de sa tâche d'exploration moléculaire, présuppose de mettre en œuvre des techniques fiables de localisation (ou tracking) de la configuration 3D de ces interfaces afin d'y associer un modèle numérique en temps-réel. Ce tracking permettra de construire un modèle virtuel de l'objet manipulé dans son environnement protéique, visualisé conjointement avec modèle réel, ce qui augmentera et complètera ainsi la représentation atomique de l'interface tangible, par d'autres représentations moléculaires classiques (structures secondaires, ball and stick, liaisons hydrogènes, champs électrostatiques...) essentielles en biologie structurale. Comme il ne s'agit de pas, dans ce travail, d'adresser des problématiques de visualisation moléculaire, nous choisirons un outil parmi ceux existants [16], pour visualiser le modèle virtuel, en le connectant au tracking de l'interface tangible. Par ailleurs, pour présenter à l'utilisateur des scores biologiquement pertinent et pour évaluer la qualité de la solution durant la construction de son modèle physique de peptide, différentes énergies seront présentées visuellement, calculées en temps réels par des outils existants de simulation moléculaire [17] [11], développés par des équipes avec lesquelles le groupe VENISE noue déjà des collaborations. Selon le contexte de recherche ou d'enseignement, le but du jeu sérieux sera par exemple d'atteindre le score maximal, correspondant à une solution de repliement ou de docking optimal proposée par l'utilisateur. Nous envisageons aussi de rendre les composants de l'interface actifs, en la connectant au jeu et en y ajoutant des feedbacks visuels lumineux ou vibro-tactiles à l'aide de leds ou de vibrateurs directement insérés dans l'interface, permettant de percevoir, durant la manipulation, des retours d'information situées.

Dans la *faraday Discussion 169*, deux types de jeux scientifiques sérieux ont été identifiés. Le premier type de jeu sérieux concerne la médiation scientifique auprès du grand public pour initier à une problématique scientifique, qui peut cependant être aussi utilisé comme cours introductif. C'est le cas par exemple de *UDock* [4], utilisé au *CNAM* sur le docking protéine/protéine, et à l'Université d'Utrecht au Pays Bas, par *Alexandre Bonvin*, pour l'enseignement du docking protéine/protéine. L'objectif du second type de jeu sérieux, plus difficile à atteindre, est de transmettre des concepts, voire de produire de la connaissance dans le domaine scientifique ciblé. C'est le cas par exemple du jeu sérieux *Fold'It* [5], dédié au repliement des protéines (folding) qui a permis l'identification de structures biomoléculaires complexes par ses joueurs (protéase du VIH). C'est ce second objectif que nous suivons dans notre approche.

La difficulté pour aboutir à ce jeu sérieux réside essentiellement dans le fait que ce type de tracking pose plusieurs problèmes difficiles : l'interface tangible est une chaîne articulée de morceaux d'acides aminés différents dont il faudra détecter la nature, la chaîne articulée sera de taille variable en fonction des constructions de l'utilisateur, les mains de l'utilisateur durant la manipulation occasionneront nécessairement une discontinuité forte dans le tracking, qu'il faudra résoudre, notamment à l'aide des contraintes d'angle connu de la chaîne polypeptidique.

Le tracking 3D d'objets rigides ou poly-articulés (corps humain, main humaine, bras robotique) a fait et fait encore l'objet de recherches intenses [13], car l'estimation de la configuration spatiale d'un objet à partir d'une observation non intrusive est par nature un problème très complexe. L'approche la plus classique consiste à utiliser un système de marquage explicite, par exemple *via* des marqueurs planaires prédéfinis (système *ARToolkit* [14]) ou des marqueurs infrarouges rétro-réfléchissants (systèmes commercialisés, par exemple *ARTTrack*). Un ou plusieurs capteurs optiques suffisent alors à calculer la localisation et la configuration spatiale de l'objet étudié. Cette approche a d'ailleurs été exploitée dans les travaux d'*Olson et al.* en biologie moléculaire [9,10].

Fiable et précise, cette approche souffre cependant de son caractère intrusif qui en limite l'utilisation dans un contexte opérationnel réaliste. Par ailleurs, d'autres contraintes intrinsèques liées à la visibilité des marqueurs fondent la recherche d'approches alternatives, sans marqueurs (markerless tracking). Du fait de la connaissance a priori sur les interfaces manipulées (types de molécules, degrés de libertés contraints entre les groupes d'atomes les constituant), une approche à base de modèle est la plus prometteuse, et a déjà montré son efficacité pour le tracking d'objets rigides [15] ou poly-articulés [18]. Les travaux les plus récents cherchent à associer plusieurs sources d'informations complémentaires : à l'extraction et la mise en correspondance robuste de contours locaux, on associe des indices colorimétriques [22], ainsi que l'information de profondeur issue de capteurs de type *Kinect*, dont l'usage connaît un développement rapide [19, 20].

A notre connaissance, ces approches récentes, développées pour le tracking d'humains et d'objets dans des environnements d'intérieur, n'ont pas encore été adaptées au suivi d'interfaces tangibles déformables. Ce type de tracking pose cependant plusieurs challenges : l'interface étudiée est une chaîne articulée de différents morceaux d'acides aminés dont il faudra détecter la nature, et dont la taille sera variable en fonction des constructions de l'utilisateur. Le tracking d'un objet modulable constituera à ce titre une innovation majeure. Enfin plusieurs axes de recherche seront investigués afin de pallier les occultations du modèle par les mains de l'utilisateur : prédiction/vérification de configuration à l'aide de contraintes d'angle connues de la chaîne polypeptidique, utilisation de capteurs embarqués dans le modèle (ex : centrale inertielle incluse dans le modèle durant la fabrication [21]).

Programme de recherche envisagé

Ce projet doctoral sera divisé en quatre grandes phases : la construction par impression 3D des briques de base du peptide à l'échelle gros grain servant d'interface tangible déformable, en s'inspirant des modèles proposées dans [5] (voir Figure 2), à l'aide du *FabLab* récemment inauguré dans le cadre de l'Équipement d'Excellence *Digiscope*. La seconde phase du projet consistera à identifier et implémenter les techniques de tracking non invasives décrites précédemment, en combinant de techniques de *markerless tracking* et de capteurs sur l'interface tangible déformable. La dernière phase du projet sera la conception, l'implémentation, en intégrant les outils de visualisation et de simulations moléculaires ciblées et du système de tracking, associé à l'identification de scénarios de difficultés croissantes, allant de la construction d'un peptide à partir des briques de bases, en passant par son repliement, jusqu'au docking

protéine/peptide. Cette phase sera suivie d'une évaluation du jeu sérieux dans un contexte d'abord de recherche sur un groupe constitué d'experts, puis dans un contexte d'enseignement.

Adéquation avec l'équipe d'accueil

Mikael Trellet, en fin de 2^{ème} année de thèse dans l'équipe VENISE, est expert en docking protéine/peptide, ayant passé deux ans à Utrecht dans l'équipe d'*Alexandre Bonvin* sur cette thématique sur le logiciel HADDOCK [7][11]. Il sera un atout important dans la réussite de ce projet, en particulier sur la validation de la pertinence biologique de l'approche.

Plus largement, l'équipe VENISE a déjà une longue expérience de conception d'interfaces multimodales pour la bioinformatique structurale, l'une des applications du projet ANR CoRSAIRE porté par P. Bourdot, concernant le docking protéine/protéine. Autrement, celui-ci, a aussi mené des recherches en reconstruction 3D en analyse et synthèse d'image, et le tracking de modèles physiques déformables de peptides, s'inscrit tout à fait dans cette thématique de recherche. D'ailleurs, il a été relecteur de la thèse d'*Alexandre Gillet* sur les premières interfaces tangibles en réalité augmentée pour la biologie structurale [9,10].

Enfin, comme l'objectif du jeu sérieux cible l'enseignement des sciences, nous nous appuyerons sur l'expérience réussie du projet ANR EVEILS [6] dont l'équipe VENISE au LIMSI fut partenaire, dont le but était de donner à percevoir dans un contexte de Réalité Virtuelle certains concepts de la relativité restreinte, afin de les transmettre de manière plus efficace aux étudiants [6]. Jean-Marc Vézien, qui sera l'encadrant scientifique principal de ce projet doctoral, fut co-leader du projet EVEILS. Sa thématique principale est la Réalité Mixte et Augmentée, et il maîtrise donc toutes les connaissances et compétences sous-jacentes en tracking par analyse d'image, nécessaires à l'encadrement et à la réussite de ce projet.

Remarques sur l'axe stratégique de rattachement

Nous insistons sur l'axe principal choisi, *Les Sciences de la Vie à leurs interfaces*, étant donné que ce travail vise des retombées importantes en biologie structurale, avec une collaboration historique de l'équipe VENISE utilisant des outils de visualisation et de simulations numériques existants [16][17][11] développés chez nos collaborateurs. Ce travail a évidemment aussi des résonances nombreuses dans l'axe « Science et innovation », à la fois par son approche de jeu sérieux, basée sur une interface tangible déformable qui sera conçue et imprimée dans le *FabLab* de l'Université, mais aussi parce qu'elle est susceptible d'avoir un impact important en terme de méthodologies innovantes, de recherche, d'enseignement, et de rayonnement de l'IDEX et plus largement de l'Université Paris Saclay.

Références

1. <http://www.rsc.org/ConferencesAndEvents/RSCConferences/FD/FD169/>
2. A.-E. Molza, N. Férey, M. Czjzek, E. Le Rumeur, J.-F. Hubert, A. Tek, B. Laurent, M. Baaden, O. Delalande, *Innovative interactive flexible docking method for multi-scale reconstruction elucidates dystrophin molecular assembly*, Faraday Discussion 169, 2014.
3. P. Chakraborty and R. N. Zuckermann, *Coarse-grained, foldable, physical model of the polypeptide chain*, PNAS, 110(33), pp 13368–13373, 2013.
4. G. Levieux, G. Tiger, S. Mader, J.-F. Zagury, S. Natkin, M. Montes, *Udock, the interactive docking entertainment system*, Faraday Discussion 169, 2014
5. S. Cooper, F. Khatib, A. Treuille, J. Barbero, J. Lee, M. Beenen, A. Leaver-Fay, D. Baker, Z. Popović and Foldit players. *Predicting protein structures with a multiplayer online game*. In Nature 466, 756-760 , 2010.
6. C. de Hosson, I. Kermen and E. Parizot, *Exploring students' understanding of reference frames and time in Galilean and special relativity*, European Journal of Physics, 31, 1527-1538, 2010.
7. M. Trellet, A. S. J. Melquiond, A. M. J. J. Bonvin, *A Unified Conformational Selection and Induced Fit Approach to Protein-Peptide Docking*, PloS one 8(3), 2013.
8. N. Férey, G. Bouyer, C. Martin, P. Bourdot, J. Nelson, J.-M. Burkhardt, *User needs analysis to design a 3D multimodal protein-docking interface*, 3D User Interfaces, 125-132, 2008.
9. A. Gillet, M. Sanner, D. Stoffler, and A. Olson, *Tangible Interfaces for Structural Molecular Biology*, Structure 13, p483–491, 2005.
10. A. Gillet, M. Sanner, D. Stoffler, and A. Olson, *Tangible Augmented Interfaces for Structural Molecular Biology*, IEEE Computer Graphics and Applications, p13-17, 2005.
11. C. Dominguez, R. Boelens, A.-M. Bonvin, *HADDOCK: a protein-protein docking approach based on biochemical or biophysical information*, J Am Chem Soc. 125(7):1731-37, 2003.
12. N. London, D. Movshovitz-Attias, O. Schueler-Furman, *The structural basis of peptide-protein binding strategies*. Structure 18:188–199, 2010.
13. V. Lepetit and P. Fua, *Monocular Model-Based 3D Tracking of Rigid Objects: A Survey*. Foundations and Trends in Computer Graphics and Vision, 2005.
14. H. Kato, M. Billinghamurst, *Marker Tracking and HMD Calibration for a video-based Augmented Reality Conferencing System*. In Proc. of the 2nd International Workshop on Augmented Reality, 1999.
15. H. Wuest, F. Wientapper, D. Stricker: *Adaptable Model-Based Tracking Using Analysis-by-Synthesis Techniques*. CAIP, 20-27, 2007.
16. Z. Lv, A. Tek, F. Da Silva, C. Empereur-mot, M. Chavent and M. Baaden, *Game on Science - how video game technology may help biologists tackle visualization challenges*, Plos one, 2013.
17. C., Yasmine, S. Pasquali and P. Derreumaux, *The Coarse-Grained OPEP Force Field for Non-Amyloid and Amyloid Proteins*, J Phys Chem B., 116(30) ; 8741-8752, 2012.
18. A. Agarwal and B. Triggs, *Tracking Articulated Motion using a Mixture of Autoregressive Models*, European Conference on Computer Vision (ECCV '04) III, pp 54--65, 2004.
19. T. Nakamura, *Real-time 3-D object tracking using Kinect sensor*. ROBIO, pp 784-788, 2011.
20. Oikonomidis, N. Kyriazis and A.A. Argyros, *Efficient model-based 3D tracking of hand articulations using Kinect*, in Proceedings of the 22nd British Machine Vision Conference, BMVC'2011, 1, 2011.
21. J. Chandaria, G. Thomas, B. Bartczak, K. Koeser, R. Koch, M. Becker, G. Bleser, D. Stricker, C. Wohlleber, M. Felsberg, F. Gustafsson, J. Hol, T. B. Schon, J. Skoglund, P. J. Slycke, S. Smeitz. *Real-Time Camera Tracking in the MATRIS Project*, IBC, Amsterdam, 2006.
22. A. Petit, E. Marchand, K. Kanani, *A robust model-based tracker combining geometrical and color edge information*, Int. Conf. on Intelligent Robots and Systems, IROS'2013, pp. 3719-3724, 2013.