

Počítačové vidění a grafika

(Josef Šivic, Jiří Matas, Michal Kozubek, Pavel Zemčík)

Cílem počítačového vidění je extrahovat informace z obrazů a videa. Automatické porozumění vizuální informaci je jedním z centrálních problémů umělé inteligence. Příklady úloh počítačového vidění jsou rozpoznávání objektů, aktivit a chování člověka, detekce lidských obličejů, sledování objektů ve videu, prohledávání vizuálních databází, nebo trojrozměrná rekonstrukce scény z několika obrazů. Počítačové vidění v současnosti zažívá průlomový vývoj díky pokroku ve strojovém učení a zejména hlubokých konvolučních neuronových sítí učených z velkého množství dat. Díky všudypřítomnosti obrazových dat a vizuálních senzorů se tak rozvíjí množství aplikací s enormním ekonomickým potenciálem. Příkladem jsou samořiditelná vozidla, kde počítačové vidění je jedním z klíčových komponent zodpovědných za porozumění scény v okolí vozidla.

Mnoho zásadních výzkumných problémů zůstává ale stále nevyřešeno a stojí v cestě k tomu, aby počítačové vidění dosáhlo nebo případně překonalo schopnosti vidění lidského. Současné metody se neumí adaptovat do nových, ještě neviděných podmínek a situací, učit se z několika málo příkladů, nebo dlouhodobě plánovat a uvažovat. Další zásadní problémy leží na cestě k automatické analýze velkého množství zašuměných dat, jako jsou například databáze složené z obrázků a textových popisů v lékařství, přírodních nebo humanitních vědách.

České počítačové vidění má ve světové komunitě silné postavení. Několik českých laboratoří počítačového vidění pravidelně publikuje mezi celosvětovou špičkou na předních konferencích (ICCV, CVPR, ECCV). Čeští vědci jsou také v organizačních výborech předních konferencí, časopisů (IJCV a TPAMI) a organizací (CVF, TCPAMI), dostali prestižní ceny (Longuet-Higgins Prize, Helmholtz Prize) a granty (ERC a ERC CZ). Jedna z hlavních Evropských letních škol počítačového vidění se už několikrát rokem koná v Praze. V roce 2004 Praha hostila ECCV - jednu z hlavních konferencí v oboru.

Počítačová grafika je oborem, který se zaměřuje na tvorbu obrazu na počítačích, přesněji na syntézu umělých obrazů (ať už realistických nebo viditelně umělých) na základě modelů objektů a scén i obrazů v počítači, a na tvorbu videosekvencí, tedy na animaci výše uvedených umělých obrazů, v řadě případů interaktivní. Rozvoj umělé inteligence a metod strojového učení, zejména konvolučních neuronových sítí, oblast počítačové grafiky značně ovlivnil. Týká se to například následujících "typických" směrů výzkumu a aplikací:

- Realistické zobrazování: Řada metod realistického zobrazování scény, jako jsou například "photon mapping" a "path tracing" vede k tvorbě výstupu, který zpočátku obsahuje malý počet vzorků (a jeví se jako zašuměný) a dále se zlepšuje (progressive refinement). Neuronové sítě se používají pro rychlé vylepšení generovaného obrazu s velmi dobrými výsledky.
- Predikce pohybu ve scéně: Postupy umělé inteligence se využívají v interakčních technikách pro predikci pohybu uživatelů na základě analýzy "historie a chování

uživatelů” a tím se zlepšují parametry interakce (reakční doba systému se zkracuje, objem přenesených dat se zmenšuje).

- Modifikace obrazu a videa: Metody umělé inteligence se v počítačové grafice využívají k řadě postupů “modifikace”. Mezi typické patří anonymizace videa (kde se záměrně “pokazí” podání lidských tváří a registračních značek aut), mezi méně typické například modifikace nápisů ve scéně (například výměna názvu restaurace při zachování fontu a stylu nápisu).
- Přenos stylů, výrazů a pohybů: Umělá inteligence se uplatňuje i v postupech, kdy je přenášén výraz tváře člověka na jiného člověka, případně na umělou tvář, kdy se syntetizuje stylizovaná podoba člověka s jiným stylem vlasů, ale i jinou rasou či zcela umělou formou tváře, ale dá se využít i k přenosu stylu pohybu člověka a/nebo zvířat i strojů.
- Aplikace ve filmovém průmyslu: Ve filmu se uplatňuje umělá inteligence v řadě postupů, jejichž příkladem může být kolorování (automatické nebo poloautomatické) černobílých filmů, podpora animace v animovaných filmech a v řadě dalších aplikací.

Zapojení pracovišť v ČR a téma výzkumu

Pracoviště	Téma
FEL ČVUT Prof. Jiří Matas, Doc. Ondřej Chum, Dr. Zuzana Kukelová, Dr. Giorgos Toliás, Doc. J. Bittner	Rozpoznávání, sledování objektů ve videu, vizuální vyhledávání, geometrické problémy, rozpoznávání a detekce obličejů a textu, fotorealistické simulátory pro generaci trénovacích dat
CIIRC ČVUT skupiny robotika a strojové vnímání (Prof. Václav Hlaváč), Aplikovaná algebra a geometrie (Doc. Tomáš Pajdla) a inteligentní strojové vnímání (Dr. Josef Šivic)	Rozpoznávání objektů a aktivit ve videu, učení ze slabě anotovaných dat, algebraické, geometrické, statistické a optimalizační metody, 3D rekonstrukce, aplikace v autonomní robotice
FI MU Brno Prof. Michal Kozubek, Doc. Pavel Matula, Doc. Petr Matula, Doc. David Svoboda, RNDr. Martin Maška, RNDr. Vladimír Ulman.	Analýza biomedicínských obrazů, zejména vývoj a benchmarking metod segmentace obrazu a sledování objektů v obraze (tracking)
FIT VUT Dr. David Bařina, Dr. Vítězslav Bera, Doc. Martin Čadík, Prof. Adam Herout, Dr. Michal Hradiš, Dr. Roman Juránek, Dr. Jan Peřiva, Dr. Marek Šolony, Dr. Michal Španěl Prof. Pavel Zemčik	Lokalizace v prostoru podle fotografií, automatická anotace snímků, kalibrace kamerových systémů, analýza dopravy (rychlost, druh a model vozidel), detekce objektů, realistické zobrazování a aplikace CNN v grafice, modifikace obrazů a videa, přesné zobrazování v CAD systémech, zpracování HDR obrazu, akcelerace algoritmů v hardware, komprese obrazu

MFF UK Doc. Jaroslav Křivánek	Metody strojového učení v počítačové grafice, simulace přenosu světla a tvorba 3D obsahu
OSU-UVA-FM Prof. Vilém Novák	Rozpoznávání obrazů, rekonstrukce signálů, automatická defektoskopie, automatické čtení poškozených znaků a symbolů, inpainting, opravy poškozených obrazů, sledování dynamických objektů.
ÚTIA AV ČR Prof. Jan Flusser, Dr. Michal Haindl, Dr. Radim Jiroušek, Doc. Filip Šroubek, Dr. Barbara Zitová, Dr. Miroslav Kárný	předzpracování obrazů, invariantní příznakový popis objektů vhodný pro jejich rozpoznávání, fúze obrazů pořízených různými senzory, segmentace i syntéza textur a fotorealistické modelování materiálů
VŠB TUO Doc. Eduard Sojka	Rozšířená realita, GPU computing, vizualizace, mobilní systémy
ZČU Prof. Josef Psutka	Lékařské aplikace, podpora zpracování, znakové řeči, podpora hledání vizuálních objektů v datech, vyhledávání informací v audiovizuálních archivech

Vybrané výsledky

- Busta M., Neumann, J., Matas J., Deep textspotter: An end-to-end trainable scene text localization and recognition framework, Proceedings of the IEEE International Conference on Computer Vision, 2017. Ocenění "Česká hlava 2018" v kategorii „technické vědy“ pro Lukáše Neumanna.
- Oquab, M., Bottou, L., Laptev, I., Sivic J., Learning and transferring mid-level image representations using convolutional neural networks. Proceedings of the IEEE conference on computer vision and pattern Recognition (CVPR), 2014. 1505 citations on Google Scholar.
- Kalal, Z., Mikolajczyk, K., Matas, J., Tracking-learning-detection, IEEE transactions on pattern analysis and machine intelligence 34 (7), 1409-1422, 2012, 2629 on Google Scholar.
- Sochor, J., Juránek, R., Špaňhel, J., Maršík, L., Široký, L., Herout, A. a Zemčík, P., Comprehensive Data Set for Automatic Single Camera Visual Speed Measurement. IEEE Transactions on Intelligent Transportation Systems. 2018.
- Ulman V, Maška M, Magnusson KEG, Ronneberger O, Haubold C, Harder N, Matula P, Matula P, Svoboda D, Radojevic M, Smal I, Rohr K, Jaldén J, Blau HM, Dzyubachyk O, Lelieveldt B, Xiao P, Li Y, Cho S, Dufour AC, Olivo-Marin J, Reyes-Aldasoro CC, Solis-Lemus JA, Bensch R, Brox T, Stegmaier J, Mikut R, Wolf S, Hamprecht FA, Esteves T, Quelhas P, Demirel Ö, Malmström L, Jug F, Tomancak P, Meijering E, Muñoz-Barrutia A, Kozubek M, Ortiz-de-Solorzano C. An objective comparison of cell-tracking algorithms. Nature Methods, vol. 14, No 12, p. 1141-1152. 2017. doi:10.1038/nmeth.4473.

- Vévoda, P., Kondapaneni, I. and Křivánek, J., Bayesian online regression for adaptive direct illumination sampling, ACM Transactions on Graphics (Proc. of SIGGRAPH 2018), 37(4), 2018.

Vybrané aplikace

- “Cell Tracking Challenge: <http://celltrackingchallenge.net/>”, a recognized international benchmark for the algorithms of cell segmentation and tracking (FI MU Brno)
- Method and an apparatus for fast convolution of signals with a one-sided exponential function, reg.: 2013, přijetí: 2016, Autoři: Seeman M., Zemčík P., Bařina D., U.S. patent US 9286268 (FIT VUT)
- Systém pro post-processing termovidea implementovaný na řipu kamery. Vyvinuto ve spolupřáci a pro společnost TESTO AG, Germany. Patent v USA No. 9386238, na trhu. (AV CR UTIA).
- Automatický systém pro řtení registrařních značek projíždějících automobilů z digitálních obrazů zaznamenaných v reálném provozu. Realizace na základě smlouvy s fy CGI IT Czech Republic, s.r.o. Systém je v provozu od r. 2017 (OSU).
- Dlouhodobá spolupřáce s firmou Toyota v oblasti autonomních automobilů od roku 2003. Roční objem v roce 2018 okolo 25 mil. Kř, podobně v roce 2019. (řVUT FEL).

Významné projekty v posledních 5 letech

- Inteligentní strojové vnímání (MřMT řR OP VVaV Excelentní řtýmy, 4.7MEur, 121MCZK, řVUT CIIRC, 2017-2022, J. řivic, CZ.02.1.01/0.0/0.0/15 003/0000468)
- Vytipování markerů, screening a řasná diagnostika nádorových onemocnění pomocí vysoce automatizovaného zpracování multidimenzionálních biomedicinských obrazů (2B06052, MřMT řR, tzv. Národní program výzkumu NPV II, 2006 - 2011, FI MU Brno, 51 MCZK)
- DISTRO – Distributed 3D object design, Marie Skłodowska-Curie ITN, No 642841 (MFF UK, J. Křivánek)
- From the cloud to the edge - smart Integration and Optimisation Technologies for highly efficient Image and Video Processing Systems, ECSEL JU - Horizon 2020, FITOPTIVIS, 2018-2021 (VUT Brno, P. Zemčík)
- ALMARVI - Algorithms, Design Methods, and Many-Core Execution Platform for Low-Power Massive Data-Rate Video and Image Processing (UTIA AV CR, Dr. Kadlec, 2014-2017)
- Centrum excellence IT4Innovations, Operařní programy EU 2011-2015 (OSU, V. Novák)
- PACMAN (Prognostics And Computer Aided Maintenance), H2020, 2016-2020 (VSB-TU E. Sojka)
- LL1303 Vyhledávání vizuálních kategorií ve velkém množství obrázků, ERC CZ, 32M CZK, 2013-2018, (řVUT FEL, O. Chum)
- Centrum pro multi-modální interpretaci dat velkého rozsahu, Centrum excellence GAřR (jediné udělené v oblasti informatiky), 2012 – 2018, Řeřitel: J. Matas, 20M CZK (pro řVUT FEL)

Vize rozvoje a přispěvku k ekonomice

- Počítačové vidění už řadu aplikací má. Příkladem jsou české startupy zabývající se analýzou dokumentů (Rossum.ai), 3D rekonstrukcí (Capturing reality, MAGIK Eye), nebo rozpoznáváním (eyedea.cz). Silná je také doména autonomního řízení, kde nadnárodní firma Valeo má v Praze vývojové centrum a Toyota společnou výzkumnou laboratoř s ČVUT FEL. Díky všudypřítomnosti obrazových dat a vizuálních senzorů je další ekonomický potenciál počítačového vidění enormní. Příklady jsou 1. doprava (autonomní vozidla), 2. bezpečnost (automatická analýza a zpracování dat z dohledových kamerových systémů), 3. zdravotnictví (nástroje automatické diagnostiky a analýzy lékařských dat), 4. zemědělství (automatické sledování zdraví plodin a hospodářských zvířat) nebo 5. zábavní průmysl (filmová postprodukce). Průlom v těchto problémech bude vyžadovat silnou interdisciplinární spolupráci počítačového vidění s dalšími obory, jako jsou strojové učení, strojové vnímání, robotika, zpracování řeči a přirozeného jazyka, teorie řízení, datové vědy, nebo automatické usuzování.
- Počítačová grafika nachází tradičně řadu aplikací v herním průmyslu (silný v ČR, příklady jsou Bohemia Interactive Studio, Amanita Design, Keen Software House či Madfinger Games), ve filmu a televizním průmyslu, v reklamě a vizualizaci (úspěšný český startup Corona Renderer), ale i v technických aplikacích. Pokrok v umělé inteligenci umožnil uplatnit v počítačové grafice nové směry výzkumu, ale zejména i nové aplikace a aplikační možnosti, z nichž jen některé byly zmíněny výše.