

VLIV CROWDINGU NA OČNÍ POHYBY PŘI SLEDOVÁNÍ VÍCE OBJEKTŮ¹

Filip Děchtěrenko (Katedra teoretické informatiky a matematické logiky, Matematicko-fyzikální fakulta UK, Česká republika) & Jiří Lukavský (Psychologický ústav AVČR, Česká republika)

Abstract. In many everyday tasks (as playing team sports or watching several children at once) people need to divide their attention among several objects. In this article we focus on Multiple Object Tracking paradigm (MOT), which is the dominant research method for the research on divided attention. In typical MOT task, several objects are presented and the goal is to track few of them for several seconds. In our research project we investigate the effect of scene complexity (number of objects) on the eye movements during the tracking task. We summarize main theories of divided attention and some known factors affecting the difficulty of the tracking task. With increasing number of objects in the scene, the tracking task becomes harder. We assume these difficulties arise from crowding. In the text we introduce the assumptions behind the project and discuss the methods used to measure the similarity of eye movement track across different experimental conditions or subjects.

Úvod

V české literatuře se informace o sledování více objektů a porovnávání očních pohybů téměř nevyskytují, rozhodli jsme tedy přispěvek koncipovat jako úvodní článek pro proniknutí do tematiky. V následujícím textu bychom chtěli uvést do problematiky distribuce pozornosti a jejího studia pomocí metody sledování více objektů. Stručně také představíme problematiku rušivých vlivů, které hrají roli při identifikaci a sledování objektů. Na závěr představíme dosavadní pozorování týkající se očních pohybů při sledování objektů.

Při každodenním životě se častokrát dostáváme do situací, kdy potřebujeme sledovat více objektů: Fotbalista potřebuje sledovat zároveň spoluhráče, protihráče a míč; Učitelka v mateřské školce musí dávat pozor na několik skupinek dětí současně; Letový dispečer pozoruje letadla přilétající na letiště. Pro úspěšné sledování více objektů je třeba pozornost.

Představa o tom, jak vypadá lidská pozornost, se měnila. Jedna z prvních teorií představila pozornost jako paprsek, který se pohyboval po zorném poli (Norman, 1968). Informace jsou zpracovávány pouze z objektů, které leží v paprsku, zatímco vše, co je mimo paprsek, zpracováváno není. Toto paradigma pracuje s konstantní šířkou paprsku, což se může jevit jako plýtvání zdrojů. Při změně pozornosti z velkého objektu na malý (např. ze slona

¹ Práce byla podpořena grantem GA ČR (P407/10/P607) „Přesuny pozornosti a oční pohyby při sledování více objektů u zdravých dospělých“.

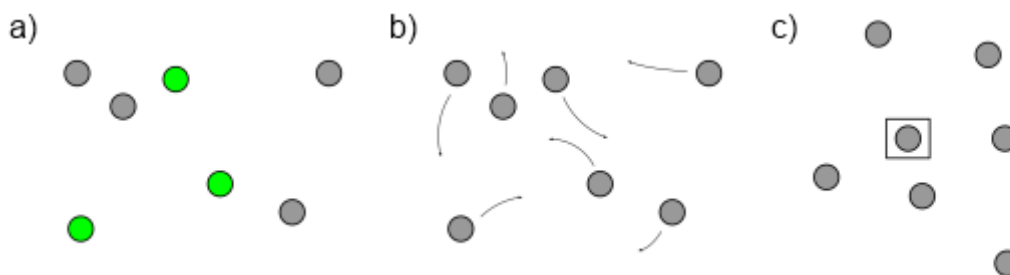
na myš) budou informace získávány i z nesouvisejících oblastí, které nebudou patřit myši. Alternativou k paprsku s konstantní šířkou může být pojetí, kdy se rozsah pozornosti přizpůsobuje zaměřenému objektu (Eriksen & St James, 1986).

Velikost oblasti, na kterou můžeme zaměřit pozornost, se může měnit, ale velikost oblasti je nepřímo úměrná kvalitě zpracování informací.

Při sledování více objektů je třeba počítat, že pozornost je obecně zaměřená na předměty na různých místech. V obou předchozích paradigmatech je pozornost vnímána jako souvislá oblast, což by bylo vážné omezení. Otázku, zda je možné rozdělit pozornost na více míst, řešili Awh a Pashler (2000). Provedli experiment, ve kterém nechali probandy zaměřit pozornost na dvě oddělená místa. Poté zobrazili písmeno buď do míst sledovaných pozorností, nebo doprostřed mezi sledovaná místa. Ukázalo se, že písmena zobrazená mezi sledovaná místa byla identifikována mnohem hůře, než písmena zobrazená do sledovaných míst. Kdyby mohla být pozornost zaměřena jen na souvislou oblast, neměla by se úspěšnost identifikace lišit. Zdá se tedy, že pozornost může být rozdělena do více míst zároveň.

Sledování více objektů

Jedno z hlavních paradigmat pro výzkum rozdělené pozornosti je Multiple Object Tracking (MOT), které zavedli Pylyshyn a Storm (1988). V typické úloze MOT je nejprve probandovi zobrazeno několik kuliček (např. 8), z nichž jsou některé označeny (obr. 1). Po nějaké době označení zmizí, kuličky se stanou nerozlišitelné a proband si musí polohu původně označených kuliček pamatovat. Poté se všechny kuličky náhodně rozpohybují po ploše. Po několika vteřinách se kuličky zastaví a proband je dotázán, zda jedna konkrétní kulička patřila do sledovaných kuliček.



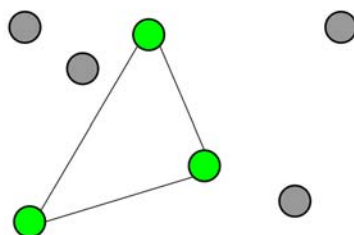
Obr. 1. Průběh Multiple Object Tracking úlohy. Nejprve je probandovi ukázáno několik kuliček (např. 8), z nichž jsou některé označeny (a). Po nějaké době označení zmizí, kuličky se stanou nerozlišitelné a proband si musí polohu původně označených kuliček pamatovat. Poté se všechny kuličky náhodně rozpohybují po ploše (b). Po několika vteřinách se kuličky zastaví a proband je dotázán, zda jedna konkrétní kulička patřila do sledovaných kuliček (c).

Úloha je vhodná na zkoumání souvislosti mezi pozorností a pohybem objektů, protože vzhledem k tomu, že sledovaných cílů je více, nelze sledovat objekty pouze pomocí fixací, ale je třeba sledovat cíle pozornostně.

Mnoho výzkumníků se věnovalo faktorům, které ovlivňují úspěšnost sledování při MOT. Průměrně zvládne člověk sledovat 4 kuličky z 8, přičemž s rostoucím počtem objektů se úspěšnost snižuje (Alvarez & Franconeri, 2007). Stejně tak snižuje výkon zvýšení rychlosti pohybu kuliček (Verstraten, 2000) či doba, po kterou kuličky sledujeme (Oksama

& Hyönä, 2004). Při zmenšení velikosti rámců, ve kterých se kuličky pohybují, taktéž dochází k snížení úspěšnosti sledování (Intriligator & Cavanagh, 2001), což se ukazuje, že souvisí s rozlišovací schopností pozornosti. Důležitým faktorem je i tvar objektu. Nejsou-li jasně odlišeny přechody mezi cíli a distraktory (např. máme-li sledovat jeden konec úsečky a druhý ne), stává se úloha obtížnější (Scholl et al., 2001). Na druhou stranu sledování neovlivňuje, zda se kuličky od sebe odrážejí, nebo zda sebou procházejí (Keane & Pylyshyn, 2006). Ukazuje se, že kuličky mohou dokonce i na chvíli mizet a sledování to neovlivní (Keane & Pylyshyn, 2006).

Bylo navrženo několik teorií vysvětlujících mechanismus sledování více cílů. Nejjednodušší teorie (switching a grouping) vychází z předpokladu, že pozorností zaměřujeme pouze jeden objekt ve vizuálním poli a operují pouze s jedním paprskem pozornosti. Switching (Pylyshyn & Storm, 1988) předpokládá, že člověk rychle ($30^\circ/\text{s} - 250^\circ/\text{s}$) pohybuje pozorností mezi jednotlivými cíli. Pozorovatel si tak aktualizuje své povědomí o poloze sledovaných objektů a předpokládá se, že by stihl navštívit všechny sledované cíle dříve, než se pohnou příliš daleko a na jejich původní místo se dostane jiný objekt. Nejjednodušší forma switchingu, která nepředpokládá ze strany pozorovatele žádnou predikci pohybu sledovaných objektů, byla vyvrácena (viz Alvarez & Cavanagh, 2005). Grouping (Yantis, 1992) vysvětluje výkon v MOT tak, že dochází k seskupení cílů do jednoho virtuálního mnohoúhelníku, který je poté opět sledován jedním paprskem pozornosti (obr. 2).



Obr. 2. Při groupingu jsou cíle seskupeny do jednoho objektu

Výzkumy ukazují (Yantis, 1992), že seskupování cílů do větších objektů pomáhá sledování.

Mezi složitější modely patří multifocal attention a FINST. Přístup Multifocal attention (Cavanagh & Alvarez, 2005) předpokládá, že naši pozornost umíme rozdělit mezi více míst. Model pracuje s několika paprsky pozornosti, které se pohybují společně s cíli. Na konci pohybu má člověk zaměřenou pozornost na stejné kuličky jako na začátku pohybu. Teorie FINST (Pylyshyn & Storm, 1988) zavádí tzv. předpozornostní fázi, ve které jsou cílům přiřazeny virtuální „prsty“ (FINgers of INSTantiation). Tyto prsty fungují jako značky, které umožňují vizuálnímu systému vybrat jeden z takto označených objektů automaticky a přesunout k němu zrakovou pozornost, která se v tomto pojetí zaměří jen na jeden objekt.

kdyby byla uspořádána horizontálně. V dolní či horní části zorného pole je situace obrácená (Feng et al., 2007).

V horní části zorného pole je efekt crowdingu silnější než ve spodní (He et al., 2000), což by mohlo mít biologické důvody – lidstvu během jeho vývoje hrozilo nebezpečí mnohem více ze země než ze vzduchu, a proto je naše schopnost rozlišovat ve spodní části zorného pole lepší.

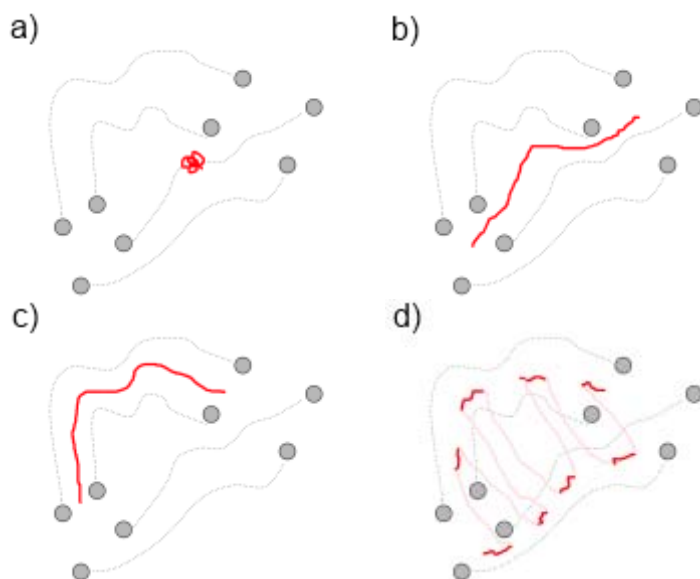
Teorií vysvětlujících crowding je mnoho (viz např. Levi, 2008) – od biologických, souvisejících se stavbou oka, až po teorie dotýkající se lidské pozornosti. Všechny ale pracují s dvojfázovým modelem. V první fázi dochází k detekci rysů objektů – zrakový systém rozpoznává barvy, tvary, jas, apod. V druhé části dochází k integraci rysů. Zorné pole je pokryto integračními poli, jež se liší svojí velikostí. Blíže fovee jsou menší a směrem k periférii se zvětšují. Nacházejí-li se dva objekty ve stejném integračním poli, mísí se rysy objektů do sebe a dochází ke crowdingu.

Ke crowdingu může docházet i při MOT. Budeme-li mít v úloze velký počet kuliček, nastanou často situace, kdy se nějaký z cílů přiblíží na periférii jednomu z distraktorů a dojde ke crowdingu.

Oční pohyby

Oční pohyby bývají často zkoumaným aspektem zrakového vnímání. Rozlišujeme několik typů očních pohybů, z hlediska zpracování informací jsou pro nás důležité fixace a sakády na jedné straně, a plynulé sledování (smooth pursuit) na straně druhé (Schutz et al., 2011). Fixace jsou okamžiky, kdy je oko relativně v klidu a my jsme schopni získávat informace z míst, na které máme zaměřenou pozornost. Pohyby mezi jednotlivými fixacemi se nazývají sakády a na rozdíl od fixací při nich nezískáváme z pozorované scény žádné informace. I během fixací se oko trochu pohybuje (mikrosakády, tremor), identifikace fixací se provádí pomocí heuristik. Plynulý pohyb se uplatňuje v situacích, kdy oko sleduje plynule pohybující se objekt. Při tomto pohybu, který je mnohem pomalejší než sakády, informace ze sítnice zpracováváme.

Vzhledem k povaze úlohy, stál výzkum očních pohybů při MOT spíše v pozadí. Pylyshyn a Storm (1988) dokonce považovali oční pohyby za nežádoucí a pokud se proband nedíval do středu rámce, byl pokus vyřazen. V poslední době se očními pohyby zabývá stále více výzkumníků. Přestože jsou cíle při MOT sledovány pozornostně, ukazuje se, že se strategie sledování projevují na očních pohybech. Lidé nejčastěji používají tyto strategie (Fehd & Seiffert, 2008, viz obr. 5): pohled do středu rámce (cíle jsou sledovány pouze pozornostně), pohled do geometrického středu cílů, souběžný pohled s jedním z cílů (ostatní cíle jsou sledovány pozornostně) a přebíhání mezi několika místy.



Obr. 4. Přehled strategií: a) pohled do středu, b) sledování centroidu, c) sledování centroidu části cílů, d) přebíhání mez skupinkami cílů

Tyto strategie je možné vyjádřit analyticky, ale žádná z nich plně nevysvětluje veškerou variabilitu očních pohybů.

Trajektorie očních pohybů v čase se nazývá scanpath. Tento termín pochází od Notona a Starka (1971) a souvisí s jejich teorií očních pohybů a zrakové paměti. Jejich teorie již dnes není přijímaná, nicméně termín scanpath se rozšířil a používá dodnes (alternativou a synonymem je termín scan pattern).

Z technického hlediska je problém, jak komplexní data, jako je sled očních pohybů, mezi sebou srovnávat. Při návrhu metriky je třeba řešit, za jak podobné či odlišné považujeme oční pohyby, které se liší polohou fixací, ale také jejich trváním či pořadím.

Pro srovnání očních trajektorií existuje několik používaných postupů. Populární metodou je metrika založená na editační vzdálenosti. Jednotlivým částem vizuální scény (např. obrazovka rozdělená na 5x5 polí) jsou přiřazena písmena, a tak lze každý záznam zakódovat jako řetězec znaků. Při definovaných cenách za změnu písmen, vkládání a mazání v řetězcích lze podobnost dvou řetězců vyjádřit jako cenu operací potřebných ke změně jednoho na druhý. Další metody jsou založené na shlukové analýze nebo korelacích. V našem případě jsme použili Normalized Scanpath Saliency (NSS) techniku popsanou Petersem et al. (2005) pro statické scény a upravenou Dorrem et al. (2010) pro dynamické scény. Tato metoda umožňuje vyjádřit míru vzájemné podobnosti uvnitř skupiny trajektorií očí a tak měřit vnitřní konsistenci tohoto chování při opakovaných prezentacích nebo porovnávat naměřené trajektorie s modelem.

Závěr

Pro výzkum rozdělené pozornosti jsou často používané úlohy MOT. Ačkoli jsou na nich nejčastěji zkoumány dílčí otázky o vztahu pozornosti a cílů, v poslední době se zvyšuje zájem o studium očních pohybů. Zatím není zřejmé, jak se projevuje sledování více objek-

tů na očních pohybech. Oční pohyby by mohly být prediktorem strategií používaných pro sledování, dokonce by mohly napovídat, zda člověk během sledování neztratil nějaký objekt. Nicméně se jedná pouze o domněnky a k jejich potvrzení budou třeba další studie.

Bibliografie

- ALVAREZ, G. A. & CAVANAGH, P. (2005). Independent resources for attentional tracking in the left and right visual hemifields. *Psychological Science* 16 (8):637–643.
- ALVAREZ, G. A. & FRANCONERI, S. L. (2007). How many objects can you track? Evidence for a resource-limited attentive tracking mechanism. *Journal of Vision* 7 (13).
- AWH, E. & PASHLER, H. (2000). Evidence for split attentional foci. *Journal of experimental psychology. Human perception and performance* 26 (2):834–846.
- BOUMA, H. (1973). Visual interference in the parafoveal recognition of initial and final letters of words. *Vision Research* 13 (4):767–782.
- BOUMA, H. (1970). Interaction effects in parafoveal letter recognition. *Nature* 226:177–178.
- CAVANAGH, P. & ALVAREZ, G. (2005). Tracking multiple targets with multifocal attention. *Trends in Cognitive Sciences* 9 (7):349–354.
- DORR, M. & MARTINETZ, T. & GEGENFURTNER, K. R. & BARTH, E. (2010). Variability of eye movements when viewing dynamic natural scenes. *Journal of Vision* 10 (10).
- EHLERS, H. (1936). V: The movements of the eyes during reading. *Acta Ophthalmologica* 14 (1–2):56–63.
- ERIKSEN, C. W. & ST JAMES, J. D. (1986). Visual attention within and around the field of focal attention: a zoom lens model. *Percept Psychophys* 40 (4):225–240.
- FEHD, H. & SEIFFERT, A. (2008). Eye movements during multiple object tracking: Where do participants look? *Cognition* 108 (1):201–209.
- FENG, C. & JIANG, Y. & HE, S. (2007). Horizontal and vertical asymmetry in visual spatial crowding effects. *Journal of Vision* 7 (2).
- INTRILIGATOR, J. & CAVANAGH, P. (2001). The spatial resolution of visual attention. *Cognitive Psychology* 43 (3):171–216.
- KEANE, B. P. & PYLYSHYN, Z. W. (2006). Is motion extrapolation employed in multiple object tracking? Tracking as a low-level, non-predictive function. *Cognitive Psychology* 52 (4):346–368.
- KOOI, F. & TOET, A. & TRIPATHY, S. & LEVI, D. (1994). The effect of similarity and duration on spatial interaction in peripheral vision. *Spatial Vision* 8 (2):255–279.
- LEVI, D. & SONG, S. & PELLI, D. (2007). Amblyopic reading is crowded. *Journal of Vision* 7 (2).
- LEVI, D. M. (2008). Crowding – an essential bottleneck for object recognition: a mini-review. *Vision research* 48 (5):635–654.
- NORMAN, D. (1968). Toward a theory of memory and attention. *Psychological Review* 75 (6), 522.
- NOTON, D. & STARK, L. (1971). Scanpaths in eye movements during pattern perception. *Science* 171 (968):308–311.

- OKSAMA, L. & HYÖNÄ, J. (2004). Is multiple object tracking carried out automatically by an early vision mechanism independent of higher-order cognition? An individual difference approach. *Visual Cognition* 11 (5):631–671.
- PELLI, D. & PALOMARES, M. & MAJAJ, N. (2004). Crowding is unlike ordinary masking: Distinguishing feature integration from detection. *Journal of Vision* 4 (12).
- PETERS, R. & IYER, A. & ITTI, L. & KOCH, C. (2005). Components of bottom-up gaze allocation in natural images. *Vision Research* 45 (18):2397–2416.
- PYLYSHYN, Z. W. & STORM, R. W. (1988). Tracking multiple independent targets: evidence for a parallel tracking mechanism. *Spatial Vision* 3 (3):179–197.
- SCHOLL, B. J. & PYLYSHYN, Z. W. & FELDMAN, J. (2001). What is a visual object? Evidence from target merging in multiple object tracking. *Cognition* 80 (1–2):159–177.
- SCHUTZ, A. C. & BRAUN, D. I. & GEGENFURTNER, K. R. (2011). Eye movements and perception: A selective review. *Journal of Vision* 11 (5).
- VERSTRATEN, F. & CAVANAGH, P. & LABIANCA, A. (2000). Limits of attentive tracking reveal temporal properties of attention. *Vision Research* 40 (26):3651–3664.
- YANTIS, S. (1992). Multielement visual tracking: attention and perceptual organization. *Cognitive Psychology* 24 (3):295–340.