

# Model pre sledovanie objektu s prediktívnym riadením zdrojov streamingu v multikamerových systémoch

*Object Tracking Model with Predictive Control of Streaming Sources in Multicameras Systems*

Dávid Cymbalák<sup>1</sup>, Miroslav Michalko<sup>1</sup>, František Jakab<sup>1</sup>

<sup>1</sup> Laboratórium počítačových sietí, Katedra počítačov a informatiky,  
Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach  
Letná 9, 042 00 Košice, Slovenská republika

david.cymbalak@cni.sk

**Abstrakt:** Práca sa zaoberá analýzou, porovnaním a návrhom prepojenia streamingových, detekčných a sledovacích technológií v prostredí multikamerových systémov. Cieľom je návrh prediktívneho sledovacieho systému, ktorý v reálnom čase sleduje a trasuje objekt v záberoch z viacerých kamier a realizuje streaming videa práve z jedného zdroja, v ktorom je objekt najlepšie zobrazený. Na ohodnotenie jednotlivých zdrojov systému bola navrhnutá metrika hodnotenia založená na aktuálnej polohe, veľkosti a miere spoľahlivosti detekcie sledovaného objektu v zábere. V práci boli navrhnuté mechanizmy a spôsoby predikcie hodnotenia jednotlivých zdrojov, predikcie nasledujúceho vysielaného zdroja a spôsoby riadenia streamingu.

**Kľúčové slová:** Video streaming, Sledovanie objektu, Počítačové videnie, Multikamerové systémy, Predikcia

**Abstract:** This work deals with the analysis, comparison and design of interconnection between streaming, detection and tracking technologies in the multi-camera environment. The goal is design of predictive tracking system with ability to track object in real time across multiple cameras and also with ability to stream video from one source with the best display of object. There was designed a metric for rating each source of system, which is based on current position, size or detection reliability of tracked object in the scene. It was also designed mechanisms and methods for prediction of video source rating, for prediction next streamed source and methods for stream control.

**Keywords:** Computer vision, Multicameras systems, Prediction methods for stream control.

## 1 Úvod

Problematika sledovania a trasovania objektu v záberoch z priestoru zachytávanom viacerými kamerami je v posledných rokoch skúmaným odvetvím najmä bezpečnostných systémov. Problémy vznikajúce pri detekcii objektu v rôznorodých záberoch rôznej farebnosti a z rôznych uhlov boli predmetom viacerých výskumov. Dôraz sa však doposiaľ veľmi nekládol na prepojenie detekčných technológií so streamingovými technológiami pre účely výstupného vysielania v reálnom čase obsahujúceho aktuálny najlepší záber sledovaného objektu v rámci multikamerového systému.

Problém výberu najlepšieho záberu na určitý objekt je možné rozdeliť na niekoľko pohľadov: z pohľadu ostrosti záberu a kvalitatívnych zložiek prenášaného obrazu, z pohľadu veľkosti objektu vzhľadom na záber, z pohľadu natočenia a uhla pohľadu na objektu alebo z pohľadu pozície objektu vzhľadom na záber. Ak je žiaduce preniesť vždy iba najlepší záber na objekt v reálnom čase s automatickým výberom záberov, je nutné uvažovať aj o probléme oneskorenia medzi prenosom obrazu a reálnou polohou sledovaného objektu. Z tohto dôvodu prichádza do úvahy predikcia výberu najlepšieho záberu a to na báze predpovede nasledujúceho hodnotenia, predpovede trajektórie pohybu alebo kombinácie so zahrnutím informácií o vzájomnej polohe kamier a konjunkcie záberov.

Cieľom práce prezentovanej v článku je teda vytvorenie prediktívneho sledovacieho systému, ktorý by na základe periodického ohodnocovania všetkých zdrojov multikamerového systému realizoval streaming práve jedného zdroja videa, ktorý bude mať v aktuálnom čase najvyššie predikované ohodnotenie na základe novovytvorenej metriky. Idea je navrhnúť takú metriku, ktorá určí také ohodnotenie jednotlivých zdrojov videa, ktoré bude závisle na aktuálnej polohe, veľkosti a ďalších vybraných vlastnostiach sledovaného objektu v zábere. Na vytvorenie návrhu jednotlivých mechanizmov navrhovaného systému je potrebné analyzovať súčasné streamingové technológie pre účely zostavenie mechanizmu riadenia streamingu, súčasné detekčné a sledovacie metódy v obraze pre účely zostavenia mechanizmu hodnotenia zdroja v multikamerovom systéme a súčasné metódy predikcie a klasifikácie pre účely zostavenia mechanizmu predikcie nasledujúceho pohybu objektu alebo nasledujúceho zdroja vysielania.

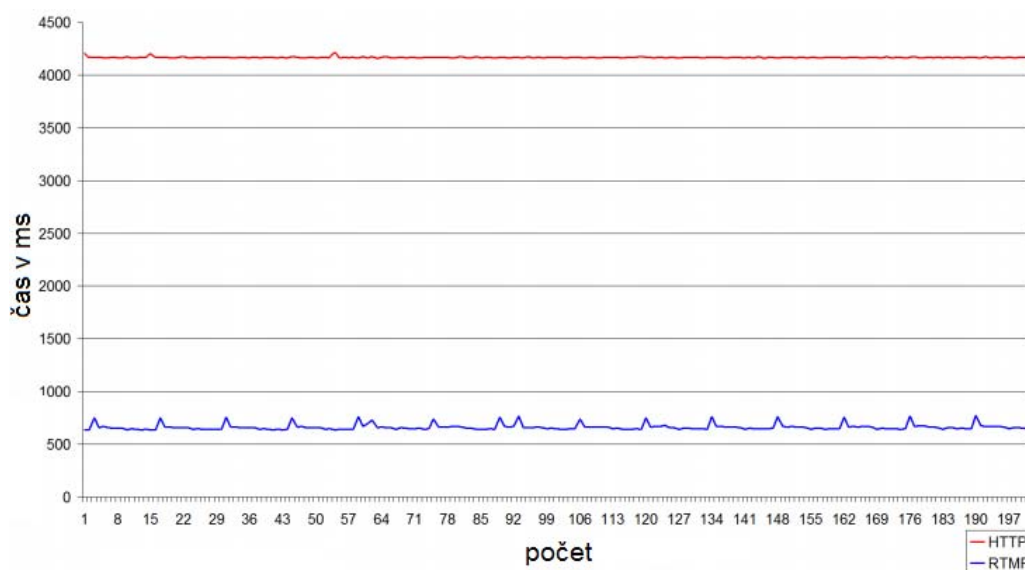
## 2 Streamingové technológie

O’driscoll (2008) definoval streamingové technológie zabezpečujú prenos multimediálneho obsahu prostredníctvom počítačovej siete vo forme vysielania zvuku, videa a doplnkového obsahu v smere od poskytovateľa k príjemcovi. V priebehu posledných rokov sa multimediálne aplikácie a streamingové technológie stali natoľko populárne, že generujú viac ako polovicu celkového internetového toku. Napriek tomu Akhshabi, Begen a Dovrolis (2011) vo svojom príspevku upozorňujú, že streamingové technológie prenášané IP sieťou sú stále realizované bez nejakej rozšírenej špeciálnej podpory zo strany siete.

Medzi neodmysliteľnú súčasť streamingových technológií patria aj streamingové protokoly, ktoré zabezpečujú prenos paketov s multimediálnym obsahom. Rozdelenie multimediálneho obsahu na viacero paketov, kompresiu a inicializáciu prenosu ku koncovému klientovi, zabezpečuje streamingový video server. Prenos multimediálneho obsahu medzi rôznymi uzlami v sieti môže byť vykonaný rôznymi spôsobmi. Typ obsahu, ktorý sa prenáša, a súvisiace sieťové podmienky zvyčajne stanovujú metódy používané pre komunikáciu. V

případe jednoduchého prenosu súboru cez siete s vyšším výskytom straty paketov je kladený dôraz hlavne na spoľahlivosť doručenia. Pakety môžu byť chránené proti stratám so zavedením redundancie alebo stratené pakety môžu byť obnovené opakovaním prenosom. Pokiaľ dochádza k doručovaniu multimediálneho obsahu v reálnom čase, najväčší dôraz je kladený na latenciu a efektívny prenos, kde môžu byť tolerované príležitostné straty. Štruktúru paketov a algoritmy pre vysielanie a riadenie prenosu v reálnom čase v sieti definujú streamingové protokoly. Podľa Begen, Akgul a Baugher (2011) sa súčasné streamingové protokoly určené na prenos multimediálneho obsahu sa líšia v implementačných detailoch, ktoré ich rozdeľujú do dvoch hlavných kategórií: protokoly s princípom tlačenia (PUSH) a protokoly s princípom ťahania (PULL). PUSH protokoly vytvoria spojenie medzi serverom a klientom, ktoré je stále udržiavané a pakety sa vysielajú ku klientovi až pokiaľ sa spojenie nepreruší vypršaním limitu alebo prerušením na strane klienta. Najznámejšie PUSH protokoly sú RTSP, RTMP, RTP, 3GPP. V PULL protokoloch je klient aktívnym prvkom, ktorý nadväzuje spojenie vo forme žiadostí o multimediálny obsah od streamingového servera. Najznámejšie využitie PULL protokolov je v prípade progresívneho sťahovania multimediálneho obsahu, kde klient posíla HTTP požiadavku na server a začína sťahovať obsah hneď ako mu server na požiadavku odpovie.

Na základe porovnania výkonnosti v súčasnosti populárnych streamingových protokolov HTTP a RTMP v prístupe realizovanom Mizerákom (2010) bol zistený rozdiel pri nasadení v trvaní spracovania požiadavky serverom. V tomto prípade bola efektivita protokolu RTMP niekoľko násobne väčšia ako HTTP (Obr. 1), avšak výsledky sa môžu líšiť v prípade nasadenia protokolu v inom type aplikácií.



Obr. 1. Porovnanie výkonnosti protokolov HTTP a RTMP v prístupe Mizeráka (2010).

### 3 Sledovanie objektu vo videu

Detekcia objektu v obraze slúži na rozpoznanie a klasifikáciu sledovaného objektu a na určenie jeho vlastností ako je veľkosť, natočenie ale aj poloha. Na rozpoznanie objektu v obraze je potrebné obraz rozdeliť na segmenty, redukovať málo významné časti obrazu,

respektíve extrahovať významné časti obrazu. Redukcia alebo extrakcia prvkov v obraze sa môže realizovať predspracovaním formou aplikácie filtrov obrazu a popisom obrazu na základe príznakov.

Z hľadiska sledovania objektu vo videu je dôležité realizovať detekciu objektu na sérii vybraných časovo odstupňovaných snímok daného videa. V každej spracovanej snímke, kde je po detekcii objekt rozpoznaný, sa vypočíta jeho poloha vzhľadom na záber snímania. Wang et al. (2011) definuje charakter pohybu objektu ako sériu údajov o polohe, či už vo forme postupnosti bodov so súradnicami alebo vo forme funkcie. Detekcia pohybujúcich sa objektov vo videu môže byť realizovaná na základe odfiltrovania pozadia. Odfiltrovanie pozadia na základe farebnosti alebo gradientu sa môže realizovať na úrovni pixelu, regiónu obrazu alebo celej snímky. Metódy, ktoré sledujú odchýlky v jednotlivých snímkach videa musia čeliť rôznym problémom, ktoré sa vyskytujú pri snímaní realistického prostredia: postupné alebo náhle zmeny osvetlenia, nízka miera farebnej odlišiteľnosti sledovaného objektu od prostredia, neželané tieň objektu, zmena pozadia sledovaného objektu alebo neustály pohyb objektu aj pri inicializácii detekcie. Všetky tieto faktory ovplyvňujú úspešnosť rozpoznania objektu vo videu a preto je vhodné použitie takzvaných učiacich algoritmov, ktoré budú priebežne sledovať zmeny na sledovanom objekte a dopĺňať modelový vzor, podľa ktorého sa objekt detekuje. Typický sledovací systém je zložený z troch komponentov: reprezentácia objektu, dynamický model a vyhľadávací mechanizmus. Objekt môže byť reprezentovaný buď holistickým deskriptorom ako je farebný histogram a hodnota úrovne jasu pixelu, alebo môže byť reprezentovaný lokálnym deskriptorom ako je lokálny histogram a informácia o farebnosti. Dynamický model je zvyčajne používaný na zjednodušenie výpočtovej zložitosti pri sledovaní objektu. Vyhľadávací mechanizmus slúži na optimalizáciu samotného sledovania objektu a môže využívať deterministické aj stochastické metódy. Dôležitým prvkom pri sledovacích metódach je pohybový model, ktorý môže vyjadrovať napríklad posuvný pohyb, transformáciu na základe podobnosti objektov alebo afinnú transformáciu.

Medzi súčasné moderné metódy používané pri sledovaní objektov v reálnom čase patria: inkrementálne vizuálne sledovanie (IVT), ktorý bližšie popisuje Ross et al. (2008), rozptýlené pomerné sledovanie (VRT) opísané Collinsom, Liuom a Leordeanu (2005), fragmentové sledovanie (FragT) opísané Adamom, Rivlinom a Shimshoniom (2006), posilnené sledovanie v reálnom čase (BoostT) popísané Grabnerom a Bischofom (2006). Sledovanie dohľadom Gong et al. (2012) možno rozdeliť na: čiastočný dohľad (SemiT) opísané Grabnerom, Leistnerom a Bischofom (2008), alebo rozšíreným dohľadom (BeSemiT) popísaným v prístupoch Staldera, Grabnera a Van Goola (2009). Ďalšie metódy ako Tracker (LIT) nasadený Meiom a Lingom (2009), sledovanie so schopnosťou viac-inštančného učenia (MIL) definované Babenkom, Yangom a Belongieom (2009), algoritmus vizuálnej sledovacej dekompozície (VTD) opísaný Kwonom a Leeom (2010) a metóda sledovania a detekcie so schopnosťou učenia (TLD) vytvorená Kalalom, Matasom a Mikolajczykom (2010), ktorá je aj podľa nasledovných výskumov Kalala, Mikolajczyka a Matasa (2012) považovaná za vysoko spoľahlivú a použiteľnú pre rozmanité typy vzorov sledovania. Spoľahlivosť a funkčnosť týchto algoritmov je možné overiť pomocou meraní úspešnosti udržania detekcie sledovaného objektu (Tab. 1) alebo pomocou sledovania chybného posunu stredovej polohy sledovaného objektu. Presnosť a spoľahlivosť sledovacieho algoritmu je možné ovplyvniť viacerými optimalizačnými metódami a úpravami podľa prístupu použitia daného algoritmu. Napríklad sledovanie tváre človeka si bude žiadať pokročilejšie techniky najmä z dôvodu zložitosti scény s viacerými tvármi, kde je podľa Zhang a Gomez (2013) nutné aby sa vzor pre sledovanie nezamieňal a neučil podobu ostatných tvári v scéne. Pri sledovaní objektu v multikamerovom prostredí je nutné uvažovať o prístupoch distribuovaného modelu sledovania, ktorý sa podľa Schrader et al. (2012) bude učiť vzor pre sledovanie z viacerých

uzlov systému. Pri sledovaní objektov, ktoré menia náhle smer alebo sa pohybujú vyššími rýchlosťami vzhľadom na záber je vhodné uvažovať o nasadení hlbokých neurónových sietí na sledovací algoritmus podobne ako definoval vo svojom prístupe Jin et al. (2013).

Vzorka videa	IVT	VRT	FragT	BoostT	SemiT	BeSemiT	L1T	MILT	VTD	TLD
Animácia 1	45	72	77	78	36	46	36	68	79	86
Animácia 2	11	8	8	8	11	11	51	8	20	70
Interiér	57	2	44	24	26	28	35	41	72	61
Šport 1	45	37	14	37	100	95	39	5	33	43
Koncert 1	56	20	21	23	27	37	100	23	99	36
Koncert 2	3	1	22	5	8	5	27	88	96	8
Šport 2	79	88	85	20	16	72	7	47	71	56
Skákavý pohyb	99	51	29	96	96	63	99	99	88	99
Auto	50	9	18	91	88	88	29	95	96	95
Otáčanie hlavy	94	3	46	85	58	51	65	89	57	81
Ľudia na ulici	21	21	3	17	66	83	24	24	14	72
Ľudia v budove	12	13	12	11	14	11	39	12	10	11
Mikročip	21	77	44	24	13	5	9	44	32	13
Animácia 3	31	10	10	26	76	91	29	26	36	57
Exteriér	12	2	16	44	12	30	42	30	40	28
Priemer	42.40	27.60	29.93	39.27	43.13	47.73	42.07	46.60	56.20	54.40

Tab. 1. Úspešnosť detekcie sledovacích algoritmov podľa Wang et al. (2011).

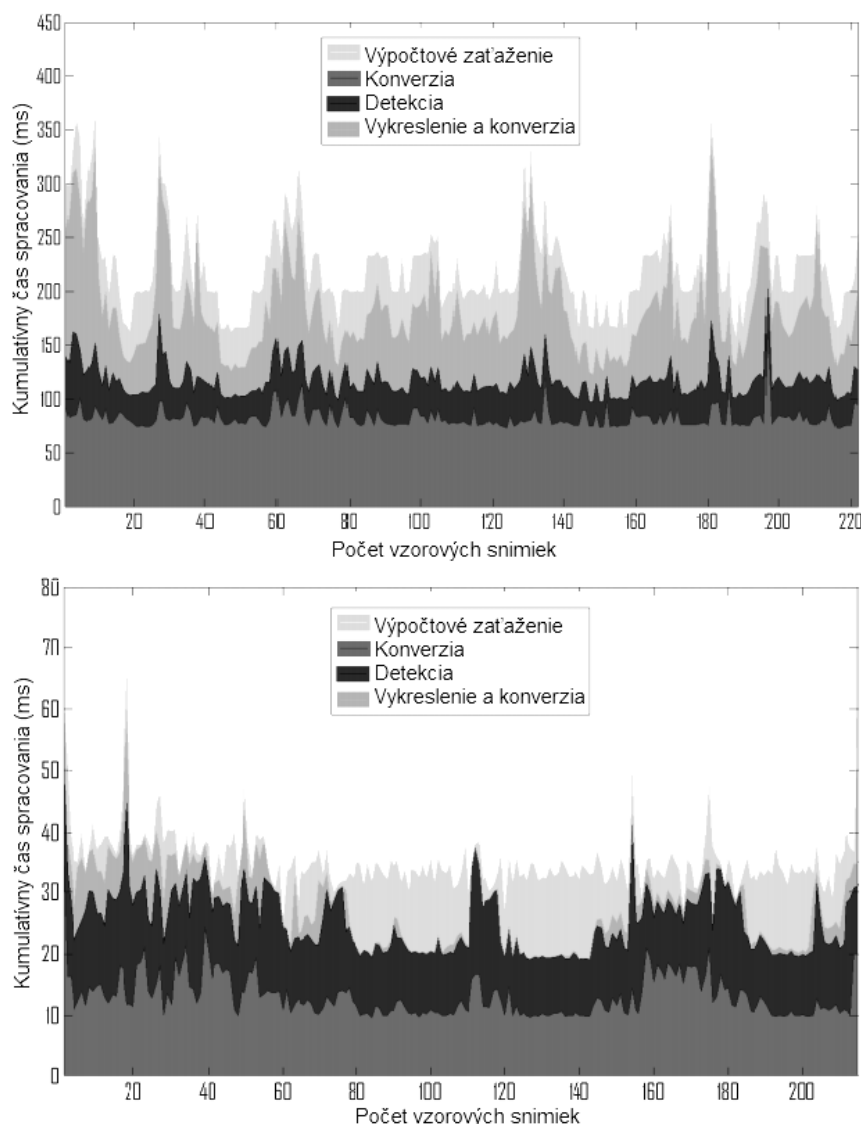
### 3.1 Knižnice počítačového videnia

#### 3.1.1 OpenCV

OpenCV (Open Source Computer Vision) je multiplatformová knižnica určená k programovaniu aplikácií počítačového videnia. OpenCV obsahuje funkcie pre spracovanie obrazu v reálnom čase, transformácie obrazu, analýzu videa, kalibráciu kamier, rekonštrukciu 3D obrazu, detekciu a sledovanie objektov, strojové učenie, multidimenzionálne hľadanie a rôzne doplnky pre tvorbu rozšírených grafických prostredí. Podľa Bradského a Kaehlera (2008) je zaujímavou možnosťou využitie výkonu grafickej karty pri realizácii počítačového videnia. V posledných verziách táto knižnica disponuje rozhraniami pre C++, C, Python a Java a je použiteľná aj v mobilných systémoch Android a iOS.

#### 3.1.2 FastCV

Knižnica počítačového videnia FastCV bola vyvíjaná firmou Qualcomm (2012), neskôr voľne prístupná. Medzi jej hlavné prednosti patria: funkcie pre spracovanie a transformáciu obrazu, detekcia, sledovanie, rozpoznávanie, 3D rekonštrukcia, funkcie pre využitie v rozšírenej realite, segmentácia, vyhľadávanie a vyspelý manažment pamäte. FastCV je primárne určené pre ARM platformy mobilných zariadení, kde dosahuje výrazne efektívnejšie a rýchlejšie spracovanie obrazu ako konkurenčné knižnice.



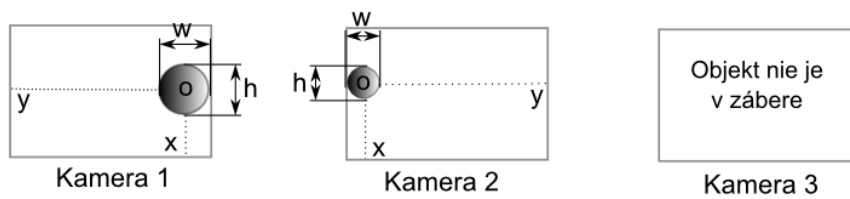
Obr. 2. Skúšobná aplikácia a zaťaženie pri OpenCV vs FastCV s prístupom od Gravdala (2012).

## 4 Návrh hodnotiacej metriky v multikamerovom systéme

Pre účely spracovania informácií o vlastnostiach sledovaného objektu v zábere kamery v čase  $t$  je dôležité poznať jeho aktuálne polohové súradnice  $x$  a  $y$  vzhľadom na záber, jeho výšku  $h$  a šírku  $w$  v zábere (Obr. 3) ale aj hodnotu dôveryhodnosti detekcie objektu  $d$ . Stav objektu  $O$  v čase  $t$  pre uzol multikamerového systému  $n$  sa môže vyjadriť ako vektor:

$$O_n^t = [x, y, w, h, d] \quad (1)$$

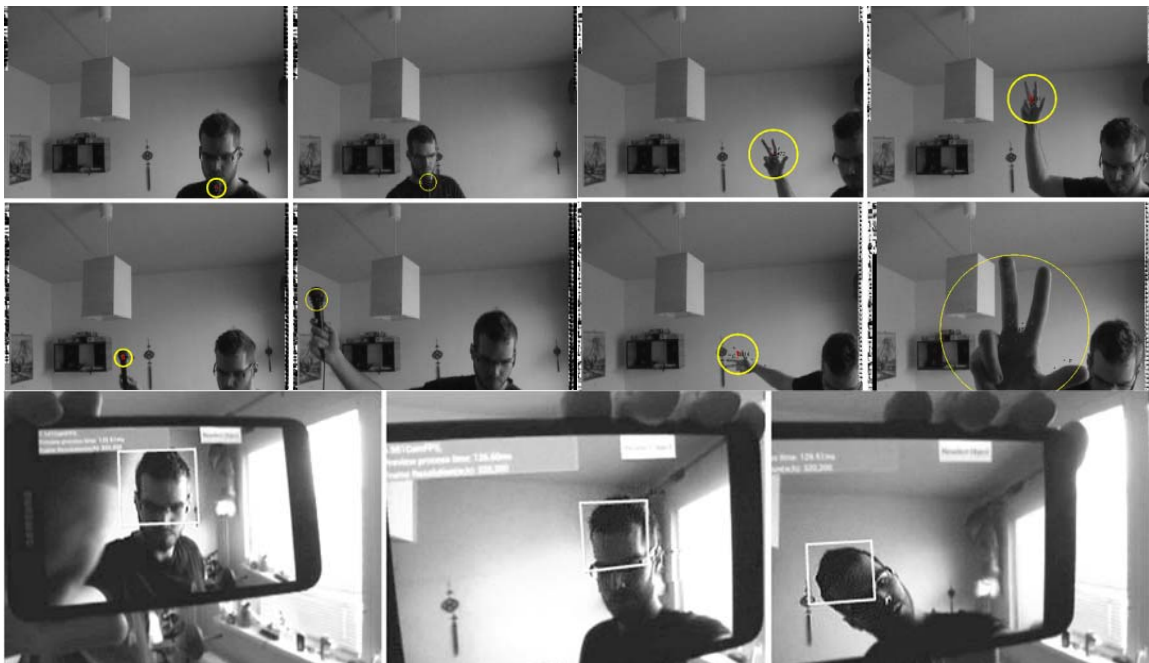
Hodnoty obsiahnuté v  $O_n^t$  budú slúžiť na výpočet metriky a porovnávanie medzi jednotlivými uzlami v multikamerovom systéme. Matica tvorená vektormi  $O_n^t, O_n^{t-1} \dots O_n^{t-k}$  bude slúžiť pre výpočet predikcie metriky v čase  $t+1$  na základe  $k$  predchádzajúcich stavov.



Obr. 3. Detekcia polohy a veľkosti objektov v čase  $t$ .

Pre získanie aktuálnych informácií o polohe a vlastnostiach objektu bol v rámci riešenia využitý algoritmus TLD a jeho implementácie openTLD. Výstup informácií o polohe objektu je možné získať priamo z výstupného súboru openTLD aplikácie, ktorý sa priebežne dopĺňa. Každý riadok výstupu má tvar  $F X Y W H D$ , kde  $F$  je poradie sledovanej snímky,  $X$  a  $Y$  je aktuálna horizontálna a vertikálna pozícia kotviaceho pixelu sledovaného objektu,  $W$  a  $H$  je pixelová šírka a výška sledovaného objektu a  $D$  je hodnota dôveryhodnosti detekcie, ktorá určuje na koľko percent je sledovaný objekt zhodný zo zdrojom. Ukážka výstupu sledovania OpenTLD C++ aplikácie s informáciami o polohe, veľkosti a dôveryhodnosti detekcie:

$F$	$X$	$Y$	$W$	$H$	$D$
489	NaN	NaN	NaN	NaN	0.000000
490	63	176	75	59	0.748394
491	64	186	75	59	0.651912
492	64	186	75	59	0.620531
493	64	184	75	59	0.613741

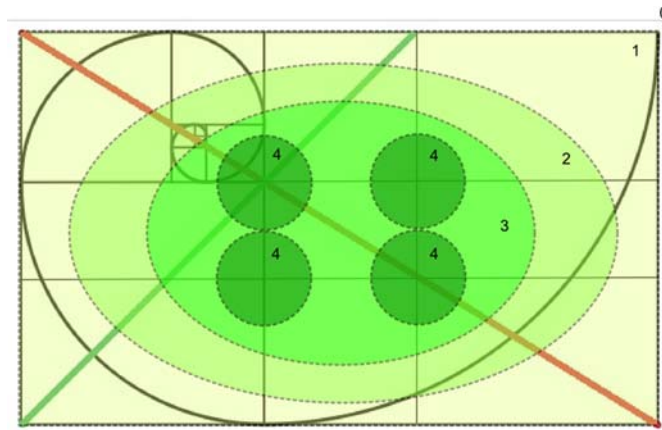


Obr. 4. Experimentálne sledovanie objektov prostredníctvom implementácie OpenTLD pre Unix a Android.

Pre účely porovnávania medzi jednotlivými uzlami systému a na ohodnotenie zdrojov videa na základe informácií o polohových vlastnostiach sledovaného objektu, je potrebné zaviesť hodnotiacu metriku, ktorá ohodnotí každý zdroj samostatne. Samotné hodnotenie by mohlo spočívať z troch zložiek: primárna zložka s polohou objektu vzhľadom na záber ( $M_p$ ),

sekundárna zložka s veľkosťou objektu vzhľadom na záber ( $M_v$ ) a terciálna zložka s dôveryhodnosťou detekcia podľa vzoru ( $M_d$ ).

Primárna polohová zložka navrhovanej metriky by mala vytvárať hodnotenie na základe aktuálnej polohy kotviaceho bodu sledovaného objektu v zábere. Záber je potrebné rozdeliť na zóny, ktoré budú reprezentovať číselné hodnotenie. Na základe aktuálnych súradníc  $x$  a  $y$  bodu objektu sa výpočtom určí, v ktorej zóne bod leží a aké hodnotenie mu prislúcha. Jednotlivé zóny môžu mať tvar sústredných elíps alebo obsahovať okrajové zóny, ktoré by určovali smer odchodu alebo príchodu objektu zo záberu. Z hľadiska optimálnej výstupnej kompozície záberu je vhodné, aby objekt ležal čo najbližšie k zlatému rezu záberu a teda rozdelenia zón bolo odvodené z Fibonacciho špirály (Obr. 5).



**Obr. 5.** Návrh zón záberu pre uzol systému s využitím optimálnej kompozície zlatého rezu.

Finálna metrika zdroja  $n$  v čase  $t$  pozostávajúca zo zložiek  $M_p$  (hodnotenie polohy),  $M_v$  (hodnotenie veľkosti),  $M_d$  (hodnotenie dôveryhodnosti) by sa mohla vyjadriť vo výslednom tvare:

$$M_n^t = (M_p^t + M_v^t) * M_d^t \quad (2)$$

Kde  $M_p^t$  je celé číslo udávajúce číslo zóny detekcie v čase  $t$ , zložka  $M_v^t$  udáva v tvare desatinného  $M_v^t \leq 1$ , pomer šírky  $w_o$  sledovaného objektu k šírke celkového záberu  $w_z$  v čase  $t$ :

$$M_v^t = \frac{w_o}{w_z} \quad (3)$$

Zložka  $M_d^t$  udáva pomer dôveryhodnosti v tvare desatinného čísla  $M_d^t \leq 1$ , ktoré definuje percentuálnu zhodnosť sledovaného objektu so šablónou vzoru určenej k detekcii. Na základe pravidelného vyhodnocovania jednotlivých zdrojov hodnotiacou metrikou sa v čase  $t$  získa vektor hodnotenia všetkých  $n$  zdrojov v systéme:

$$H_t = [M_1^t, M_2^t, \dots, M_n^t] \quad (4)$$

Po výpočte maxima  $\max_{0 \leq k \leq n} M_k^t$  výstupných hodnotení  $n$  zdrojov v čase  $t$  sa získa poradie zdroja  $k$ , ktoré určuje, ktorý zdroj videa sa má aktuálne vziať ako výstupný zdroj pre streaming videa.



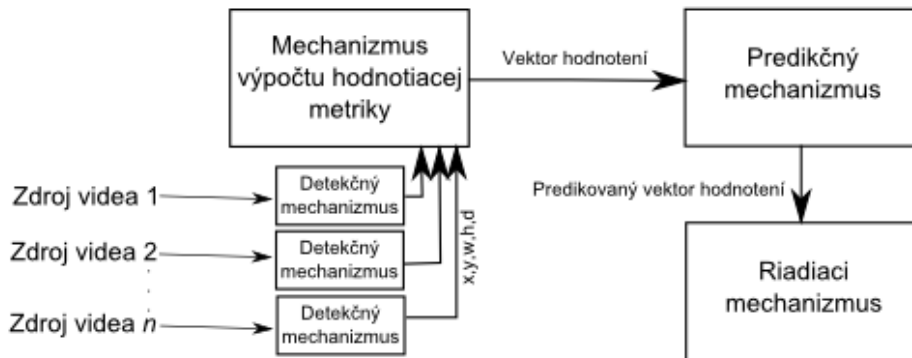
Pre optimalizáciu časového oneskorenia medzi zistením zdroja s najvyšším ohodnotením a finálnym zobrazením výstupného zdroja videa u príjemcu je vhodné zaviesť predikciu ohodnotenia, ktorá by mohla inicializovať v čase  $t$  prepnutie zdroja videa na zdroj, ktorý by mal v čase  $t+1$  získať najvyššie hodnotenie. V ideálnych podmienkach by tento prístup mohol zabezpečiť prepnutie zdroja ešte pred príchodom objektu do záberu a teda vymazať rozdiel medzi inicializáciou zmeny zdroja streamingu a skutočnou polohou sledovaného objektu v danom zábere.

Predikčné mechanizmy na odhadnutie nasledujúceho ohodnotenia zdrojov by sa do celkového návrhu mohli integrovať buď globálne a to medzi mechanizmus výpočtu hodnotiacej metriky a riadiaci mechanizmus (Obr. 6). V tomto prípade by predikčný mechanizmus vzal do úvahy  $PH_k$  maticu obsahujúcu  $H_t$  vektory  $k$  predchádzajúcich hodnotení ( $M_n^{t-k}$ ) jednotlivých  $n$  zdrojov:

$$PH_k = \begin{pmatrix} M_1^t & M_2^t & \dots & M_n^t \\ M_1^{t-1} & M_2^{t-1} & \dots & M_n^{t-1} \\ \vdots & \vdots & \ddots & \vdots \\ M_1^{t-k} & M_2^{t-k} & \dots & M_n^{t-k} \end{pmatrix} \quad (5)$$

Na základe tejto matice  $PH_k$  by globálny predikčný mechanizmus vytvoril nový predikovaný vektor hodnotení obsahujúci jednotlivé predikované hodnotenia uzlov systému pre  $n$  zdrojov v čase  $t+1$ :

$$H_{t+1} = [M_1^{t+1}, M_2^{t+1}, \dots, M_n^{t+1}] \quad (6)$$



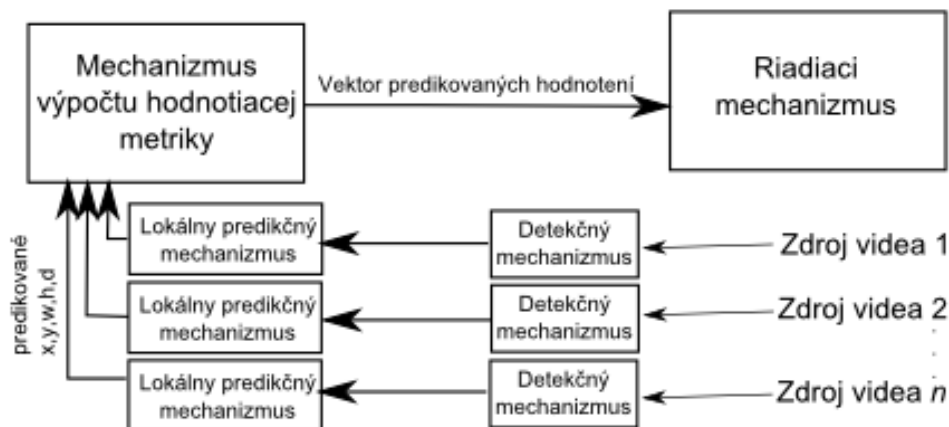
Obr. 6. Návrh umiestnenia globálneho predikčného mechanizmu.

Ďalším spôsobom je možné predikčné mechanizmy integrovať na lokálnej úrovni medzi detekčný mechanizmus a mechanizmus výpočtu hodnotiacej metriky (Obr. 7). Pri tomto spôsobe by každý lokálny predikčný mechanizmus vzal do úvahy  $PO_k$  maticu obsahujúcu  $O_n^t$  vektory  $k$  predchádzajúcich vlastností objektu (súradnice  $x$  a  $y$ , rozmery  $w$  a  $h$  a dôveryhodnosť  $d$ ):

$$PO_k = \begin{pmatrix} x^t & y^t & w^t & h^t & d^t \\ x^{t-1} & y^{t-1} & w^{t-1} & h^{t-1} & d^{t-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x^{t-k} & y^{t-k} & w^{t-k} & h^{t-k} & d^{t-k} \end{pmatrix} \quad (7)$$

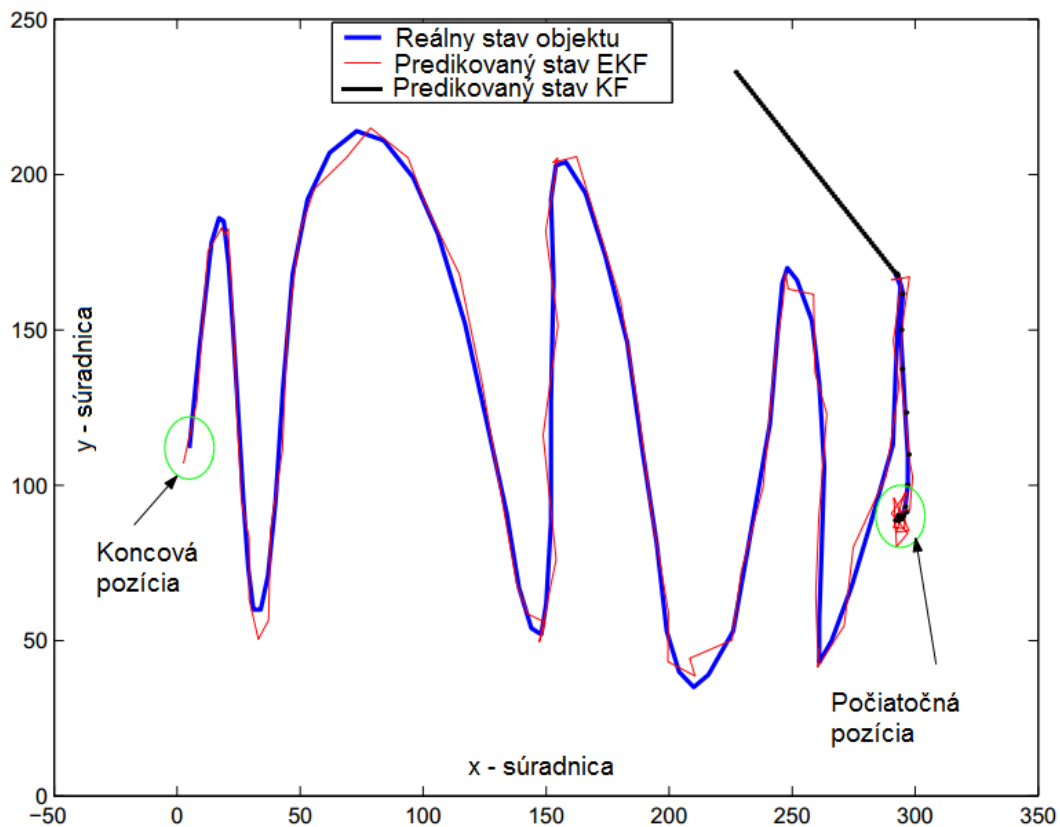
Na základe  $PO_k$  sa vytvorí lokálne pre každý zdroj  $n$  nový vektor obsahujúci predikované súradnice, veľkosť a dôveryhodnosť v čase  $t+1$ :

$$O_n^{t+1} = [x^{t+1}, y^{t+1}, w^{t+1}, h^{t+1}, d^{t+1}] \quad (8)$$



**Obr. 7.** Návrh umiestnenia lokálnych predikčných mechanizmov.

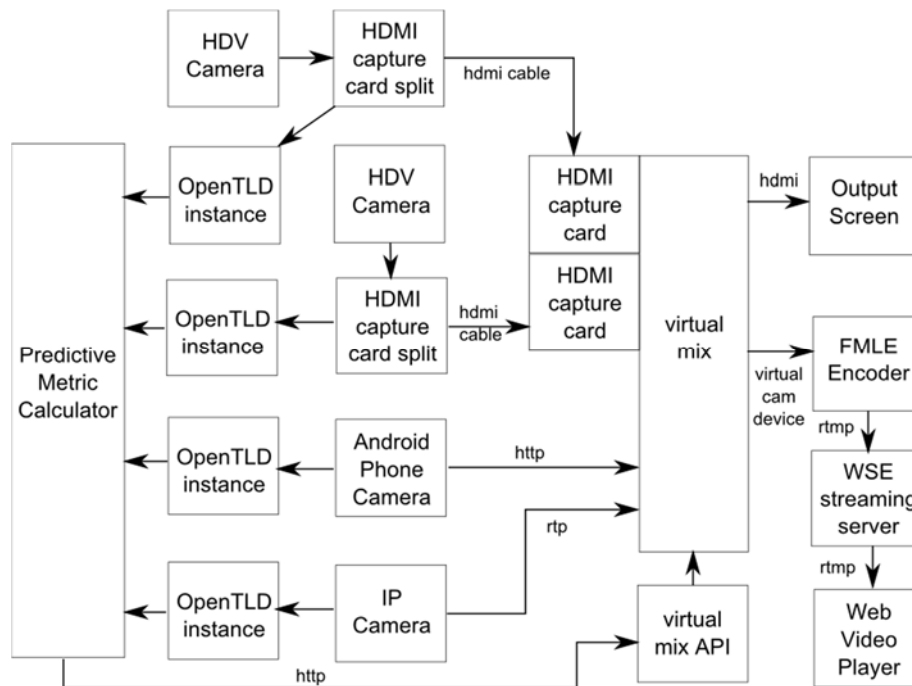
Predikčné mechanizmy na lokálnej úrovni by mohli byť založené aj na metódach predikovania trajektórie pohybu podľa predchádzajúcich súradníc ako sú Kalmanov filter alebo rozšírený Kalmanov filter. Kalmanov filter (KF) sa používa na predikovanie nasledujúcej polohy na základe opisu pohybu poslednej pohybovej sekcie. Funk (2003) definuje, že model Kalmanovho filtra potrebuje poznať reprezentáciu sledovaného objektu v stavovom ponímaní. Alternatívnym prístupom pre zvýšenie presnosti predikcie je použitie rozšíreného Kalmanovho filtra (EKF), v ktorom funkcie stavu môžu byť aj v nelineárnej podobe. Presnosť EKF je použiteľná najmä pri problematických pohyboch s náhlou zmenou smeru pohybu sledovaného objektu (Obr. 8).



**Obr. 8.** Presnosť predikcie s využitím KF a EKF pri náhlych zmenách v prístupe Cuevasa , Zaldivara a Rojasa (2005).

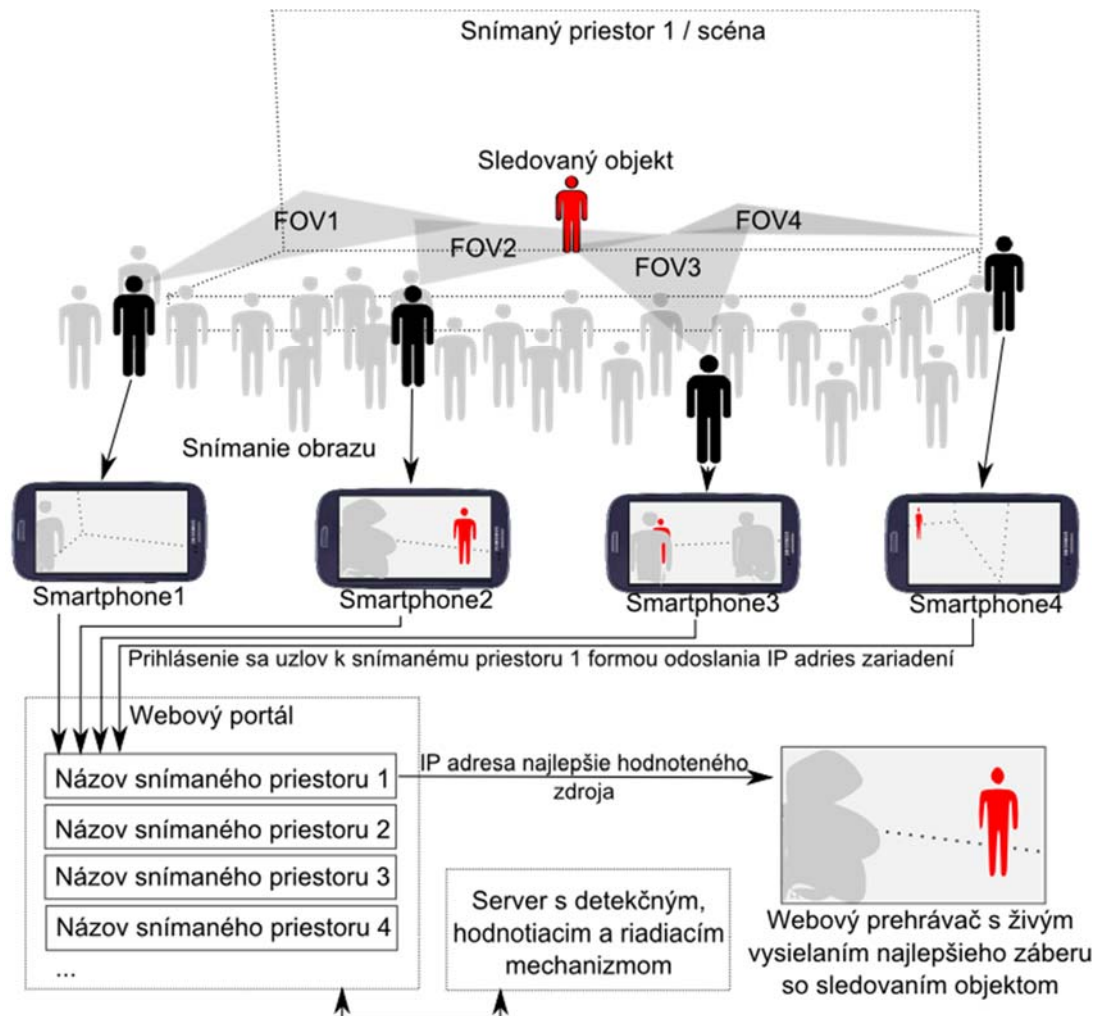
## 5 Aplikácia návrhu a možnosti využitia

V súčasnosti je navrhovaný systém v experimentálnej prevádzke a pripravený testovať zmeny v definícií metriky a v použití predikčných metód. Aplikácia riešenia pozostáva z prepojenie softvérového mixu s vytvoreným API, ktoré zabezpečuje prepínanie zdrojov streamingu v závislosti od informácií doručovaných mechanizmom pre výpočet metriky, ktorý vzchádza z lokálnych inštancií openTLD aplikácií. Táto experimentálna zostava umožňuje integrovať do systému okrem priamych káblových vstupov bežných kamier s nízkym oneskorením aj HTTP alebo RTP stream z IP kamier alebo mobilných streamingových aplikácií (Obr. 9).



**Obr. 9.** Experimentálne zapojenie systému v súčasnom stave.

Využitie navrhovaného systému by mohlo byť široké, či už pri bezpečnostných systémoch, športových podujatiach, dopravnom dispečingu, vzdelávacích konferenciách alebo nasadení na monitorovanie výrobnéj linky a podobne. Systém by mohol byť aplikovateľný na sústavu IP kamier v sieti, na mobilné snímacie zariadenia alebo na priamo zapojené lokálne snímacie zariadenia ako HD kamery a videomixážne pulty. Zaujímavou možnosťou je aplikovanie navrhovaného systému do prostredia inteligentných mobilných telefónov s integrovanou kamerou (Obr. 10).



Obr. 10. Koncepcia využitia riešenia s mobilnými zariadeniami.

## 6 Záver

Práca sa v úvode venovala analýzou súčasných technológií a prístupov v oblasti prenosu multimediálneho obsahu, detekčných a sledovacích mechanizmov pre multikamerové systémy. Cieľom tejto analýzy bolo oboznámiť sa a porovnať jednotlivé metódy, technológie a prístupy, s pomocou ktorých by bolo možné zostaviť prediktívny systém pre riadenie streamingu na základe detekcie polohy sledovaného objektu v multikamerovom prostredí. Základný návrh systému počíta s využitím viacerých alternatív aplikácie streamingových technológií s rôznou pozíciou mechanizmu riadenia streamingu v rámci kompozície multikamerového systému, streamingového servera a mechanizmu prehrávania videa. Z hľadiska detekcie sa na základe porovnania jednotlivých algoritmov javil ako najspoľahlivejší sledovací algoritmus TLD pre lokálnu detekciu objektu vo videu v rámci jednotlivých uzlov systému. Výstup tohto algoritmu obsahujúci informácie o aktuálnej polohe sledovaného objektu a ďalších je vlastností je vhodný ako vstup pre navrhnutú metriku hodnotenia jednotlivých zdrojov videa. Návrh systému taktiež predstavuje viacero spôsobov predikcie nasledujúceho hodnotenia zdroja videa, predikcie nasledujúceho vysielaného zdroja alebo predikcie očakávaného pohybu sledovaného objektu medzi jednotlivými uzlami multikamerového systému. V budúcnosti je potrebné definovať a optimalizovať metódy predikcie hodnotenia jednotlivých zdrojov, čo by malo spočívať v definovaní neurónovej

siete, respektíve Bayesovho klasifikátora a následne porovnanie spoľahlivosti predikcie v zostavenom riešení.

Ďalším zaujímavým riešením pre rozšírenie tejto problematiky by sa mohlo javiť posielanie informácií o aktuálnej polohe a ďalších vlastnostiach sledovaného objektu formou pridania dodatočných hlavičiek a úpravy vybraného streamingového protokolu. Z hľadiska nasadenia a implementácie navrhovaného systému v prostredí mobilných zariadení by bolo potrebné optimalizovať jednotlivé streamingové, detekčné a sledovacie mechanizmy tak, aby pre ich spoľahlivú funkčnosť postačoval výkon inteligentného mobilného telefónu. Je však predpoklad, že v blízkej budúcnosti dosiahnu mobilné zariadenia nových výkonových limitov, ktoré nebudú mať problém s nasadením streamingového servera priamo na zariadenie, ktoré by zároveň okrem posielania streamu objekt vo videu detekovalo, hodnotilo v reálnom čase s veľmi krátkymi intervalmi výberu posudzovanej snímky videa.

### PodĎakovanie

Tento príspevok vznikol vďaka podpore v rámci operačného programu Výskum a vývoj, pre projekt: Univerzitný vedecký park TECHNICOM pre inovačné aplikácie s podporou znalostných technológií, kód ITMS: 26220220182, spolufinancovaný zo zdrojov Európskeho fondu regionálneho rozvoja. Podporujeme výskumné aktivity na Slovensku/Projekt je spolufinancovaný zo zdrojov EÚ.

### Zoznam použitých zdrojov

- Adam A., Rivlin E., Shimshoni I. (2006). Robust fragments-based tracking using the integral histogram. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition* (pp. 798–805).
- Akhshabi S., Begen A., Dovrolis C. (2011). An Experimental Evaluation of Rate-Adaptation Algorithms in Adaptive Streaming over http. In *MMSys'11*, USA.
- Babenko B., Yang M., Belongie S. (2009). Visual tracking with online multiple instance learning. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition* (pp. 983–990).
- Bradski G., Kaehler A. (2008). *Learning OpenCV*. Sebastopol: O'Reilly.
- Begen A., Akgul T., Baugher M. (2011). *Watching Video over the Web*. Part I, Streaming Protocols, Cisco.
- Collins R., Liu T., Leordeanu M. (2005) Online selection of discriminative tracking features. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(10), 1631–1643.
- Cuevas E., Zaldivar D., Rojas R. (2005). Kalman filter for vision tracking. *FUDOCs*, Technical Report B 05-12.
- Funk N. (2003). *A Study of the Kalman Filter applied to Visual Tracking*. Project for CMPUT 652.
- Gong, C. et al. (2012). The extended co-learning framework for robust object tracking. In *Proceedings - IEEE International Conference on Multimedia and Expo* (pp. 398–403).
- Grabner H., Bischof H. (2006). On-line boosting and vision. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition* (pp. 260–267).
- Grabner H., Leistner C., Bischof H. (2008). Semi-supervised on-line boosting for robust tracking. In *Proceedings of European Conference on Computer Vision* (pp. 234–247).
- Gravdal E. (2012). *Augmented Reality and Object Tracking for Mobile Devices*. Trondheim: NTNU.
- Jin, J. et al. (2013). Tracking with deep neural networks. In *47th Annual Conference on Information Sciences and Systems, CISS 2013*.
- Kalal Z., Matas J., Mikolajczyk K. (2010). P-N learning: Bootstrapping binary classifiers by structural constraints. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition* (pp. 49–56).
- Kalal, Z., Mikolajczyk, K., Matas, J. (2012). Tracking-Learning-Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34 (7), 1409–1422.

- Kwon J., Lee K. (2010). Visual tracking decomposition. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition* (pp. 1269–1276).
- Mei X., Ling H. (2009). Robust visual tracking using l1 minimization. In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 1436–1443).
- Mizerák T. (2010). *Protokoly pro komunikaci klientu na platforme Flash*. Brno: MUNI.
- O'Driscoll G. (2008). *Next generation IPTV services and Technologies*. Hoboken: Wiley.
- Qualcomm Inc. (2012). *FastCV Library 1.1.1*. Retrieved from <https://developer.qualcomm.com/docs/fastcv/api/index.html>.
- Ross D. et al. (2008). Incremental learning for robust visual tracking. *International Journal of Computer Vision* 77(1), 125–141.
- Schrader, S. et al. (2012). A distributed online learning tracking algorithm. In *12th International Conference on Control, Automation, Robotics and Vision* (pp. 1083-1088).
- Stalder S., Grabner H., Van Gool A. (2009). Beyond semi-supervised tracking: Tracking should be as simple as detection, but not simpler than recognition. In *Proceedings of IEEE Workshop on Online Learning for Computer Vision*.
- Wang Q. et al. (2011). An Experimental Comparison of Online Object Tracking Algorithms. In *Proceedings of SPIE: Image and Signal Processing Track*.
- Zhang , T., Gomez, H.M. (2013). Technology survey on video face tracking. In *Proceedings of SPIE - The International Society for Optical Engineering*.