

3D techniky počítačového vidění

Richter Miloslav, UAMT FEKT VUT Brno

Zabývá se zpracováním signálu, především obrazu. Realizoval několik průmyslových aplikací na měření nebo detekci vad při výrobě.

Řízení HW a zpracování naměřených dat realizoval převážně pomocí programů v jazyce C/C++, který je v těchto oblastech využíván díky svým vlastnostem jako je přenositelnost, rychlost, dostupnost dat, kvalitní překlad ...

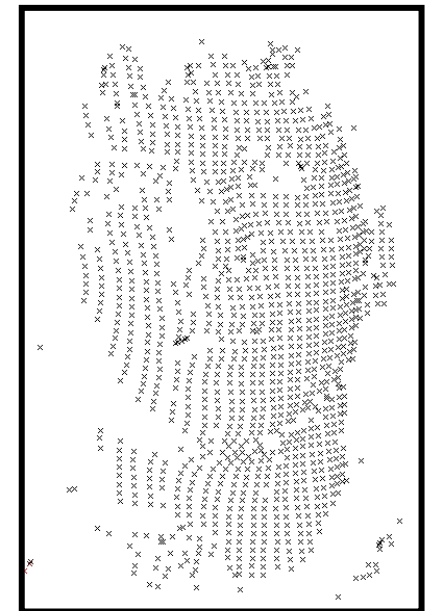
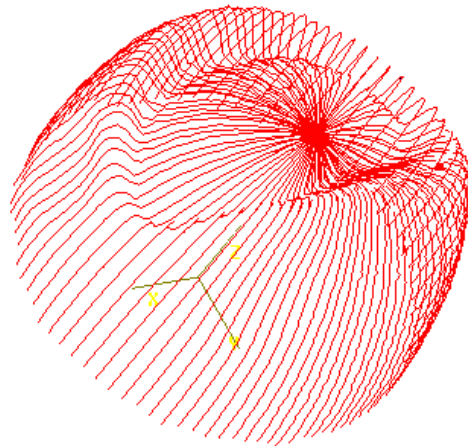
Charakteristika 3D měření

Výhody:

- bezkontaktní měření polohy, vzdálenosti, tvaru, povrchu
- zjištění/měření (prostorových) rozměrů
- zjištění 3D tvaru – “reverzní inženýring” – modely existujících věcí, modely pro virtuální realitu
- zprostředkovaně měření rychlosti (s informací o časech snímání)

Problémy:

- velké množství dat pro přenos, zpracování, archivaci
- složité (sekvenční) zpracování dat
- složitá reprezentace a vizualizace výsledných (3D) dat
-



Alternativy 3D měření

měření laserem

- doba letu, jednotlivý paprsek, nebo svazek
- výhodou je velká vzdálenost měření a dostatečná přesnost
- nevýhodou je značná doba měření
- interferenční měření (fázový posun)

radary

- doba letu
- rychlost pomocí dopplerova efektu, změna vzdálenosti za čas

(ultra)zvuk

- možnost měření i plošných objektů (pomocí pole detektorů)
- přesná metoda pro omezené množství bodů

HW pro 3D počítačové vidění

Mezi firmy dodávající průmyslové kamery a příslušenství patří Basler, PointGrey (nyní FLIR), DALSA, Matrox, Pleora, Sony, The Imaging Source, Stemmer Imaging, Allied Vision Technologies, National Instruments.

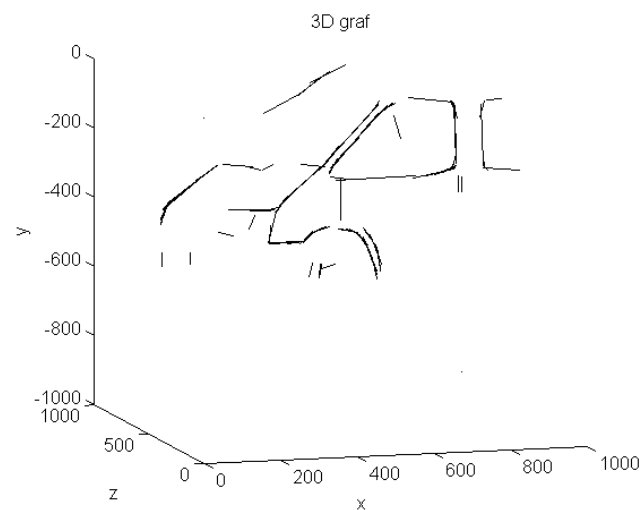
Knihovny pro práci s 3D daty (pokud nejsou dodávány pouze spustitelné aplikace):

- SDK výrobců kamer – specializace pro daný typ HW, ne vždy splní očekávání
- HALCON (<http://www.halcon.com/>)
- Matrox Imaging Library (<http://www.matrox.com/imaging/en/products/software/mil/>)
- MATLAB (<http://www.mathworks.com/products/image/>)
- OpenCV (<http://opencv.org/>)

(použité názvy jsou TM – viz odkazy)

HW standardy

- GigE Vision, USB 3.0 Vision, CoaXpress, CameraLink – "HW" normy
- EMVA 1288 – popisuje a definuje jak zjistit/změřit parametry kamery nebo detektoru – pro porovnání
- Snaha pro unifikaci a možnost vytvářet znovupoužitelné aplikace (3rd party) – GenICam, parametry driverů, optimalizace toků dat
- GenICam (Generic Interface pro kamery). Tři části: definice Api – xml soubory pro konfigurace, nastavení a řízení kamer. TL – definice bufferů a transferu dat – např. umístění a tvar RGB složek v signálu. Konvence jmen – definice datových typů, metod a jejich jmen (RoThetaPhi16, LineScan, Cartesian ...).
- normy pro výsledné 3D data – body nebo plochy. Kartézské, sférické souřadnice, hloubkové mapy a textury, určení platnosti a důvěry dat, ...

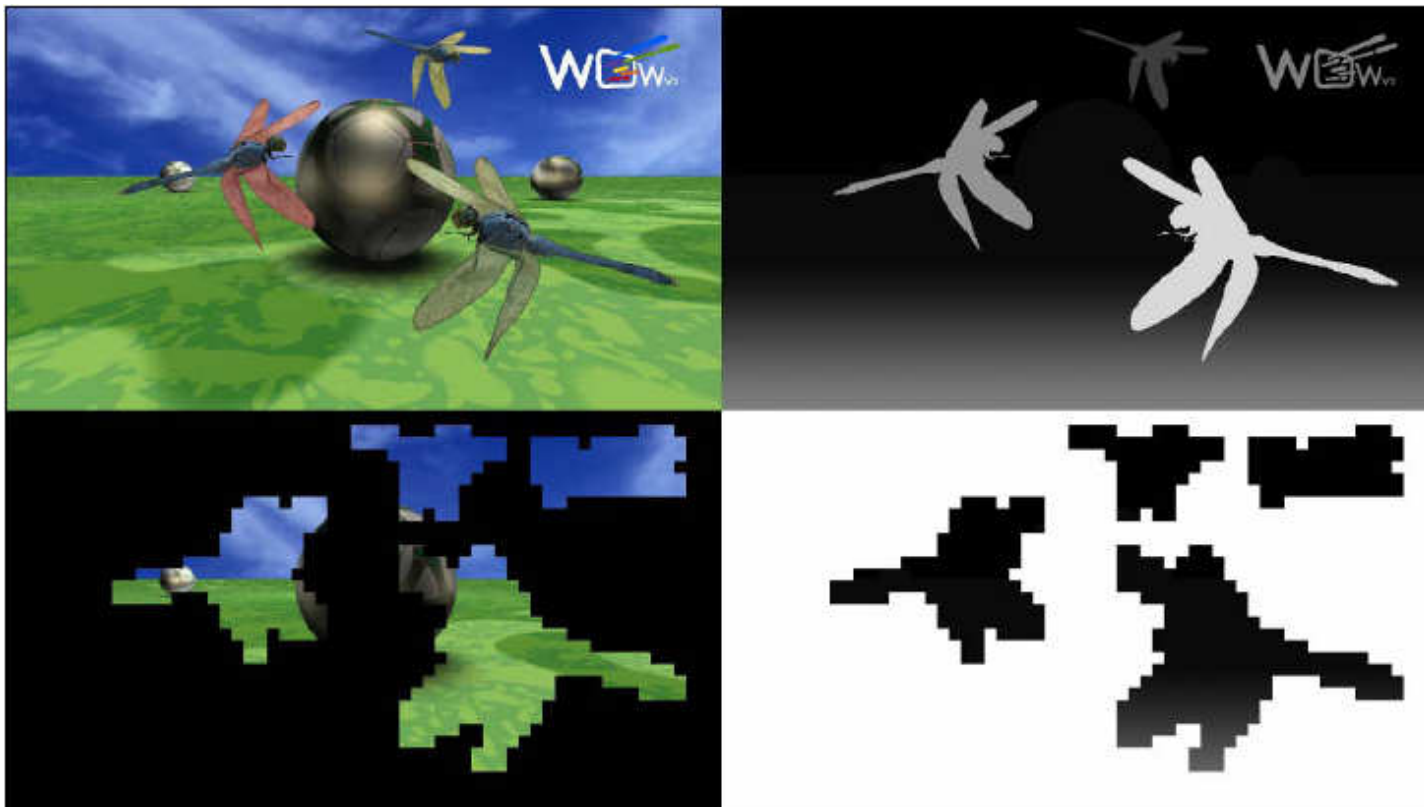


<http://www.visiononline.org/>

association for vision information – sdružení firem pro standardy týkající se součástek, kamer a vidění



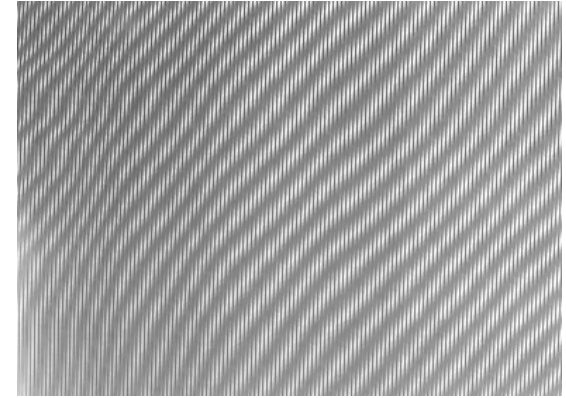
WOWvx 2D-plusDepth formát / WOWvx Declipse formát



3D metody počítačového vidění

- interferenční metody – citlivé a přesné metody, měření spojitých ploch, malý rozsah hloubky. S pomocnou mřížkou, nebo interference světelných vln.

interferenční obraz – od maxima k maximu je rozdíl jedné vlnové délky (rozdíl jeden proužek mezi mřížkami referenční a měřící)



- 3D z hloubky ostrosti – využití pro malé vzdálenosti, nutnost snímání sady snímků při přesné znalosti posunu detektoru. Snímek s nejostřejším detailem určuje polohu detailu (ve směru pohybu detektoru)
- 3D tvar ze stínu – objekt je postupně nasvécován z různých (známých) směrů. Podle poměrů jasů lze určit normálové vektory v daných místech.
- triangulace

Triangulace

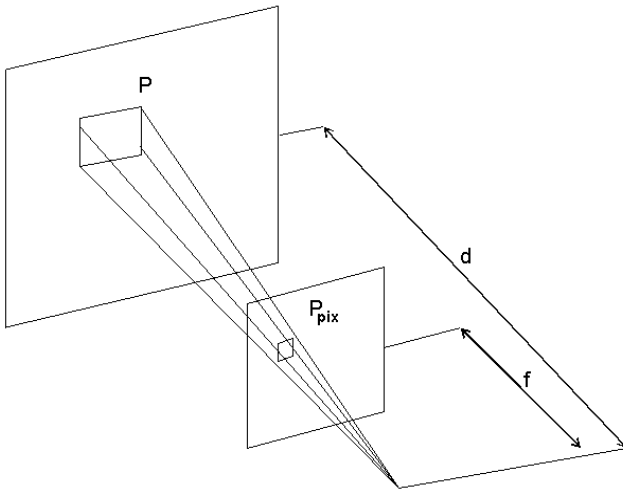
- využívá se protínání paprsků z různých směrů
- paprsek je tvořen spojnicí centra projekce a elementu detektoru (pixel) nebo paprskem světla projektoru
- základem je různost pohledů na měřený objekt
 - dáno pohybem detektoru nebo objektu



Metody triangulace

- pasivní – používají se pouze kamery
- aktivní – používá se nasvícení vzorem pro zvýšení počtu bodů, nebo přímo jako součást měřícího systému
- jedna kamera – pro 3D musí být dodána apriorní znalost (velikost objektu \leftrightarrow poloha), málo přesné
- dvě a více kamer – jsou potřebné jednoznačně definované "body" viditelné z různých kamer, body může dodat "projektor"
- jedna a více kamer ve spojení se světelným zdrojem – světelný zdroj je nutné nakalibrovat aby se mohl stát součástí měřícího systému

Z čeho plyne přesnost měření

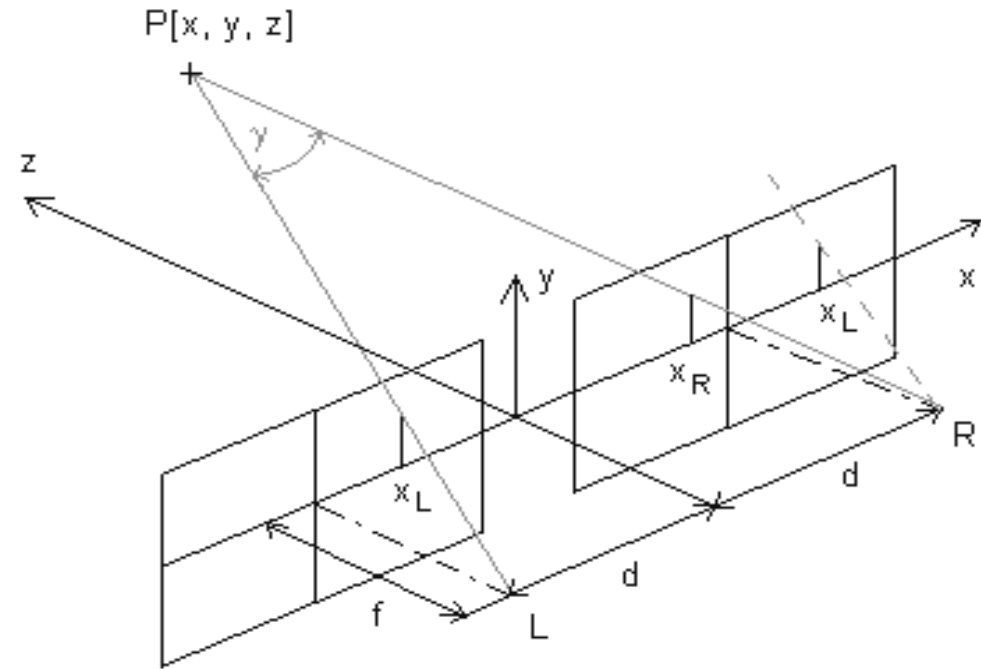


- určení polohy s tolerancí - pixel není bod, ale pokrývá určitý prostor (u dvou kamer je průsečíkem prostorový útvar)
- při zobrazení se nezobrazí bod, ale dojde k rozostření
- u dvojice kamer záleží na jejich vzájemné vzdálenosti (bázi)
- dvojice kamer musí snímat synchronizovaně u pohybujících se objektů (u horších CMOS kamer není zaručeno snímání snímku v jednom okamžiku)

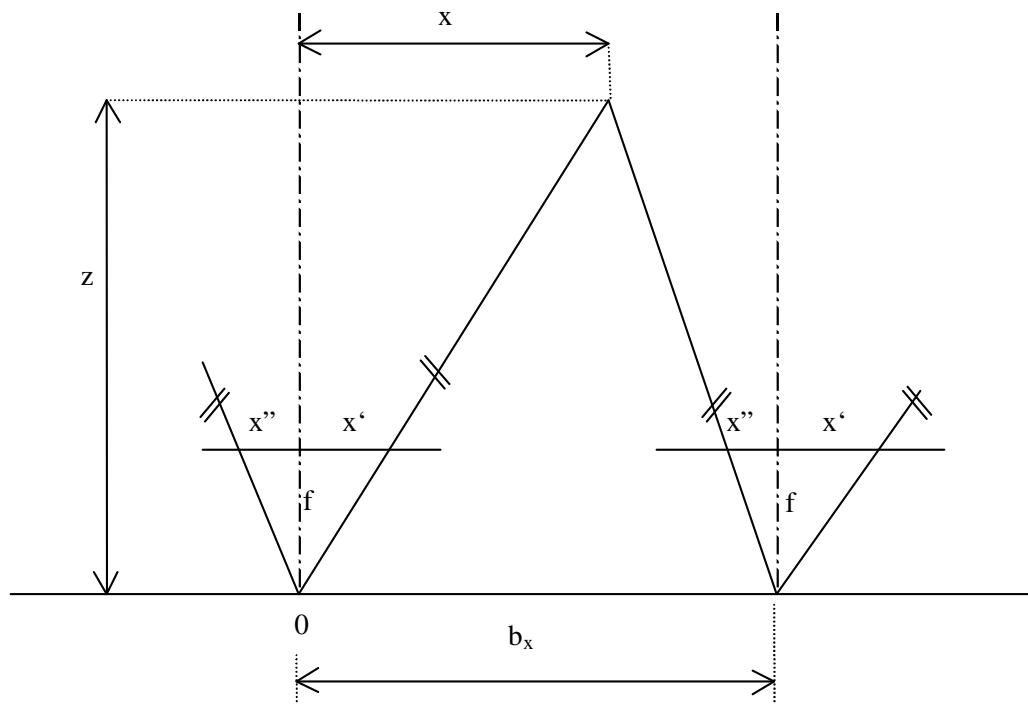
- u dvojice kamer je nutné najít společné body, bod nemusí být přítomen na všech pohledech
- uplatňuje se též šum
- vliv odrazů na povrchu (např. odraz světelného zdroje na kouli je závislý na místě pohledu)
- vliv textury (čím více je k dispozici kvalitních bodů, tím lepší je určení polohy)

Stereofotogrammetrie

- měření pomocí dvojice kamer s rovnoběžnými osami
- jednoduché vzorce,
- nevýhodou je velká chyba ve vzdálenosti díky souhlasnému pohledu
- některé body (plochy) nejsou společně vidět na obou pohledech
- chyby nastavení nejsou výrazné
- podobné vnímání lidského oka (binokulární vidění)
- využití ve snímání pro 3D vidění člověka: anaglyphy, 3D fotografie, 3D video, ...



Stereofotogrammetrické výpočty

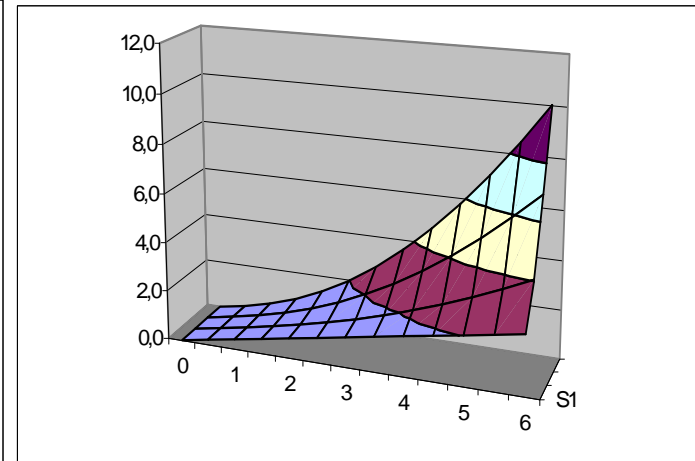
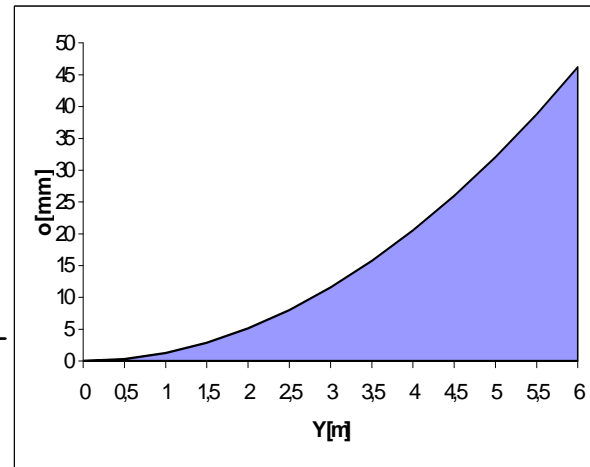
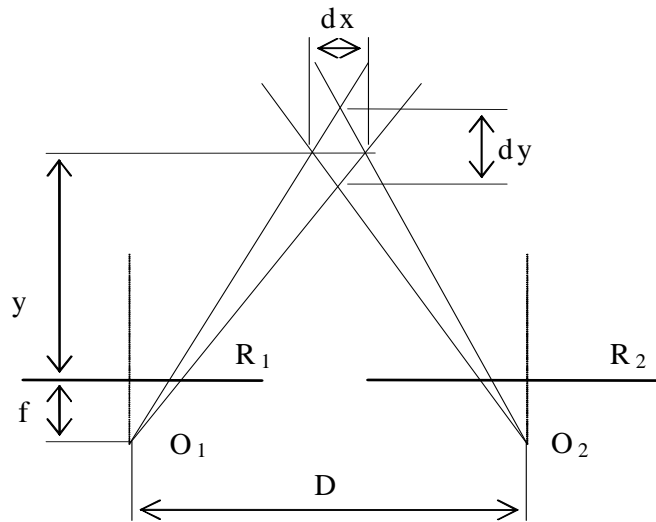


$$p = x' - x''$$

paralaxa – úhlová změna pozice (způsobená pohybem)

$$x = x' \frac{b_x}{p} \quad y = y' \frac{b_x}{p} \quad z = f \frac{b_x}{p} \quad \text{výpočet souřadnic}$$

Chyby měření způsobené diskretizací při stereofotogrammetrii

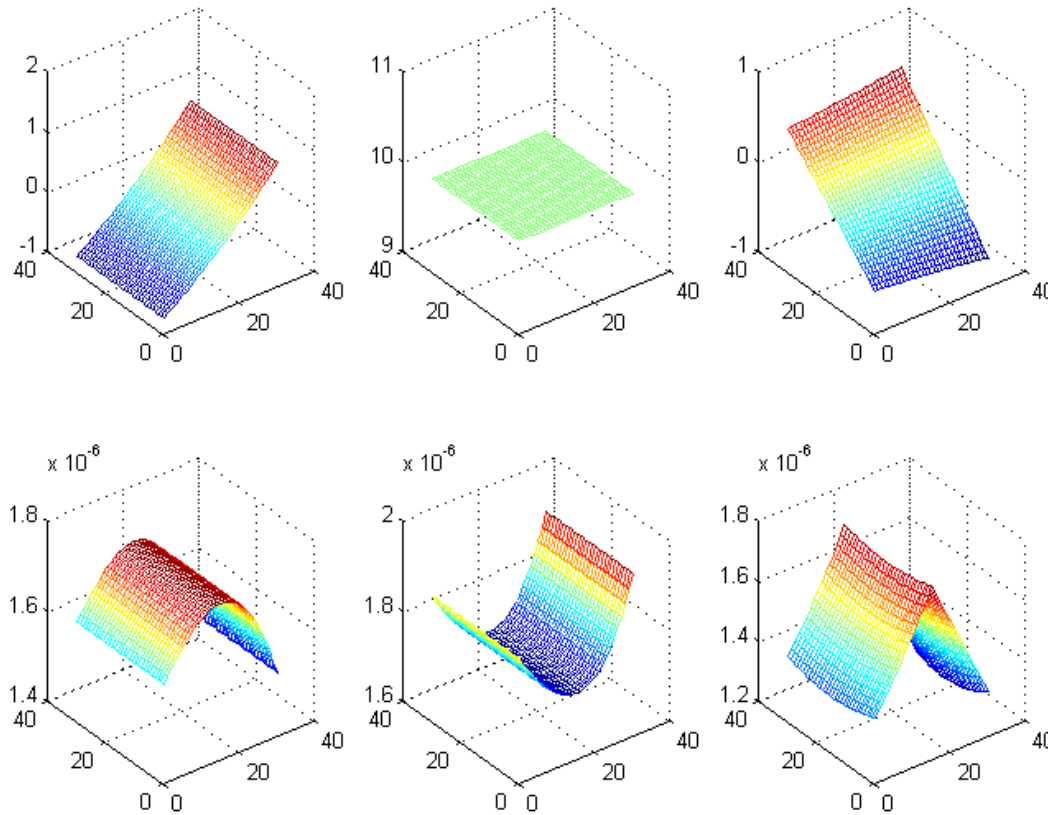


$$\sigma_Z = m \cdot \frac{Z}{D} \cdot \sigma_{xp} = \frac{Z^2}{f \cdot D} \cdot \sigma_{xp}$$

$$\sigma_X = \sqrt{\left(\frac{x_1}{f} \cdot m \cdot \frac{Z}{D} \cdot \sigma_{xp} \right)^2 + (b + \sigma_x)^2}$$

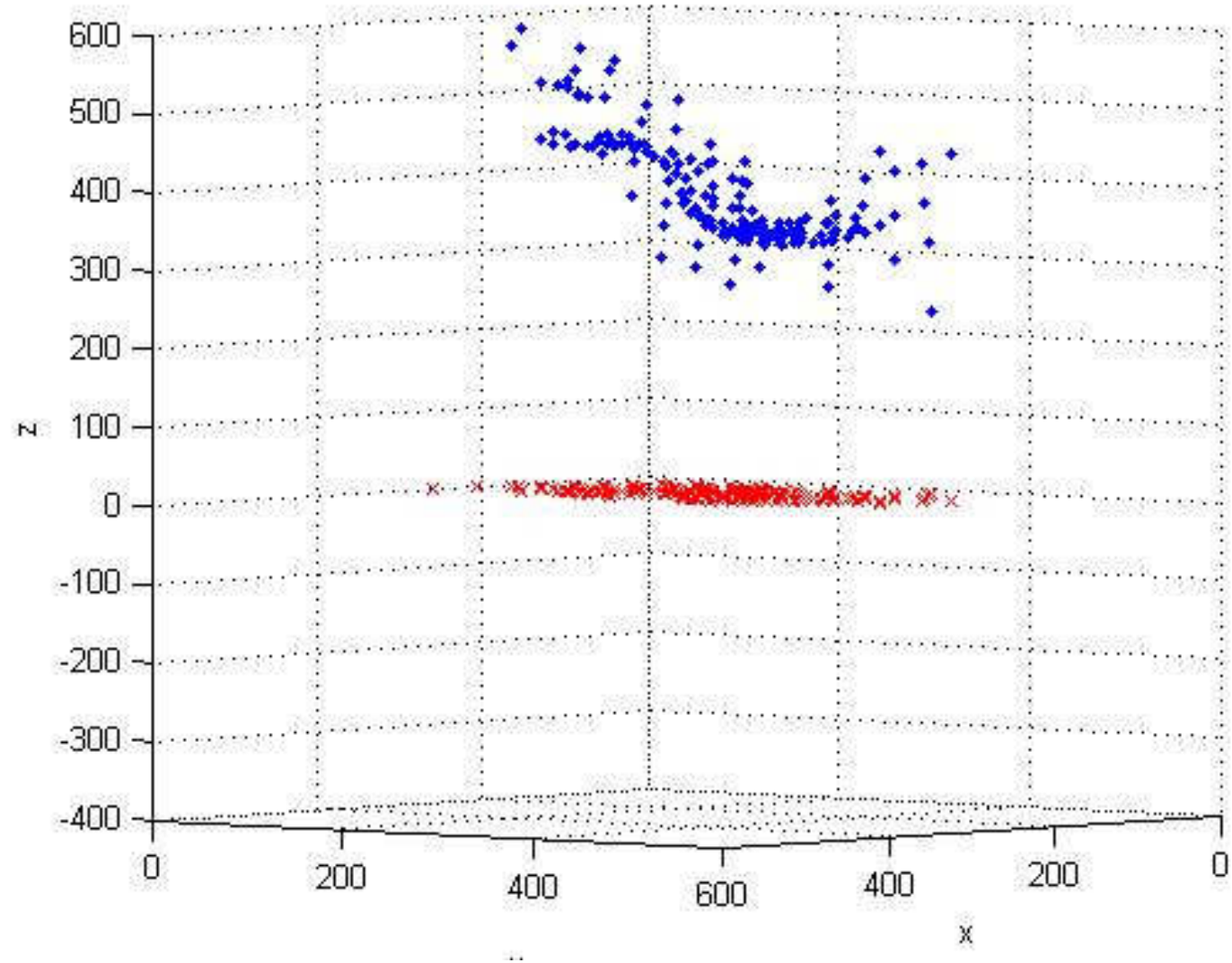
$$\sigma_p = \frac{1}{\sqrt{3}} \cdot P_{pix}$$

Měření rovinné plochy pomocí stereofotogrammetrického páru



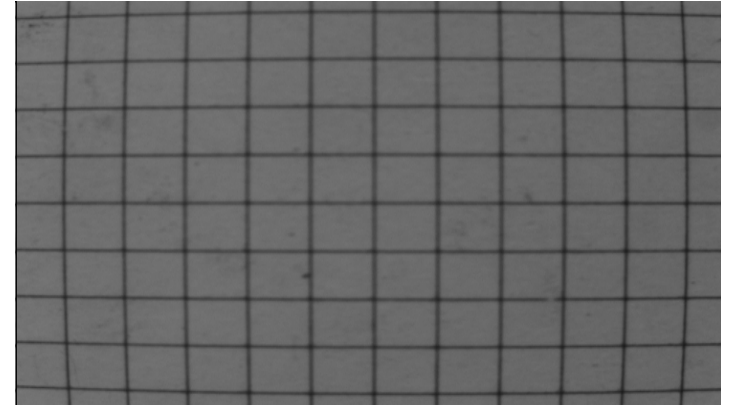
Dvě kamery, vzájemný úhel pohledu kamer 90° , dívají se do bodu na ose kamer. Horní řada udává polohu x, vzdálenost od kamery, y. Dolní řada udává příslušné chyby určení těchto hodnot.

Vliv zkreslení optiky na výsledný průběh měřené roviny



Stanovení korekčních koeficientů zkreslení

- na základě kalibračního vzoru – nejlépe 3D
- výpočet se skládá ze základního výpočtu a korekčních členů
- je možné stanovit i během výpočtu 3D souřadnic

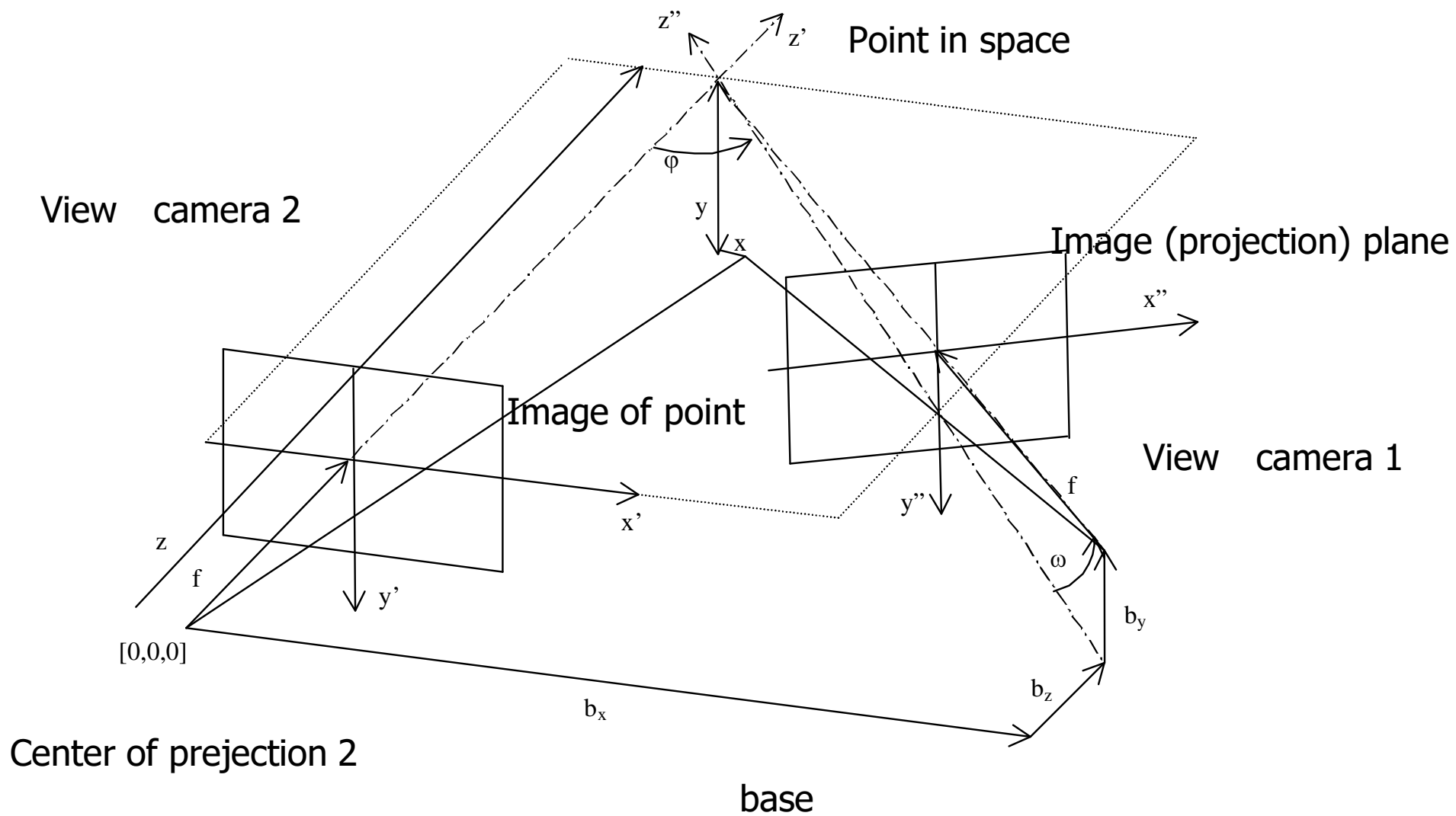


$$[x, y] = c \left[\frac{X}{Z}, \frac{Y}{Z} \right] \quad \text{centrální promítání}$$

$$\frac{dr}{r} = \frac{dc}{c} \Rightarrow dc = ca_3 r^2 + ca_5 r^4 + \dots \quad \text{radiální zkreslení}$$

$$dx = \frac{X}{Z} ca_3 r^2 + \frac{X}{Z} ca_5 r^4 + \dots \quad \dots \text{ korekce radiálního zkreslení ve směru X}$$

Obecná fotogrammetrie – sestava dvou pohledů pro 3D měření



Popis součástí sestavy obecné fotogrammetrie

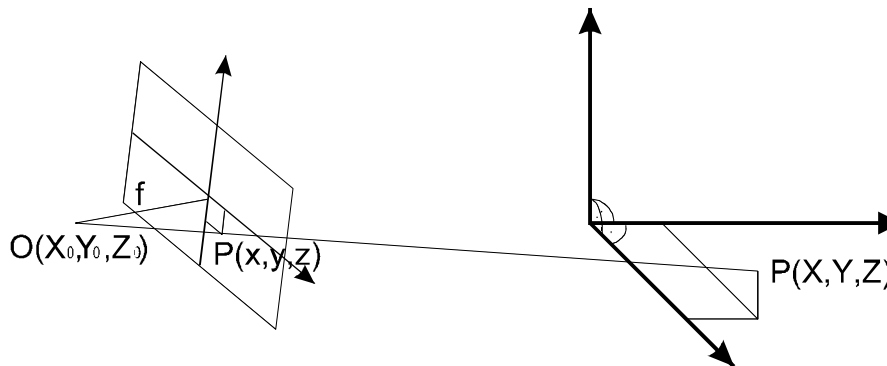
- scéna – snímaný prostor – spojený s reálnými souřadnicemi
 - snímek – výsledný 2D průmět scény pořízením snímačem s optikou (2D jasové pole)
 - pohled – udávající prostorové parametry pořízených snímků. Reprezentováno transformací souřadnic
 - kamera – určuje způsob zobrazení scény. Reprezentováno vnitřní orientací (parametry zobrazení)
 - objekty ve scéně (definované svými rozměry a jasovými vlastnostmi)
 - více kamer, objektů či pohledů spojených pevnou transformací (stereo pár, složitější objekt...)
 - pohyb (lze přidat ke všem součástem, včetně kombinací)
-
- více pohledů dává větší přesnost určení polohy – body jsou vidět z více míst, což přispívá ke zvýšení přesnosti v „dálce“ oproti stereofotogrammetrii
 - vypadne-li bod na některém ze snímků, není to problém, je-li na ostatních

Matematický aparát obecné fotogrammetrie

- rovnice koplanarity

$$Q = \sum (b \cdot v_1) x \cdot v_2$$

b – báze; v_1 a v_2 směrové vektory pohledů na daný bod z jednotlivých kamer; Q – kriteriální funkce udávající „objem“ tělesa tvořeného bází a vektory (ideálně rovina $\Rightarrow Q = 0$)



vztah mezi souřadnými soustavami

transformační rovnice

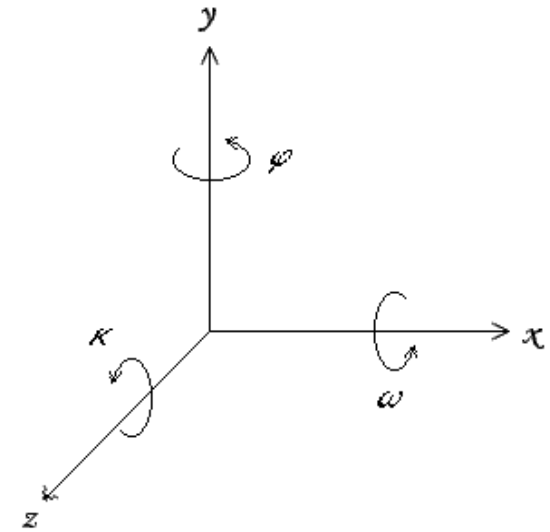
$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = R^* \left(\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} - \begin{bmatrix} X_o \\ Y_o \\ Z_o \end{bmatrix} \right)$$

složky kriteriální funkce

$$u = f \frac{r_{11}(X - X_u) + r_{12}(Y - Y_u) + r_{13}(Z - Z_u)}{r_{21}(X - X_u) + r_{22}(Y - Y_u) + r_{23}(Z - Z_u)} - u_0$$

$$v = f \frac{r_{31}(X - X_u) + r_{32}(Y - Y_u) + r_{33}(Z - Z_u)}{r_{21}(X - X_u) + r_{22}(Y - Y_u) + r_{23}(Z - Z_u)} - v_0$$

- Pro stanovení vzájemné polohy stačí 5 bodů
- obecně se doporučuje alespoň 8 bodů
- výpočet vzájemného natočení – relativní orientace – je nezávislé na měřítku, lze využít menší počet bodů
- absolutní orientace – šest stupňů volnosti (3 úhly rotace a 3 posunutí), stačí dostatek bodů a jedna vzdálenost pro změřtkování
- metoda řešení se nazývá „bundle adjustment“ a využívá minimalizace kriteriálních funkcí
- každý promítnutý bod přináší dvě rovnice (pro průmět x a y)
- alternativa: Rodrigues angles (definováno pomocí směrového vektoru)

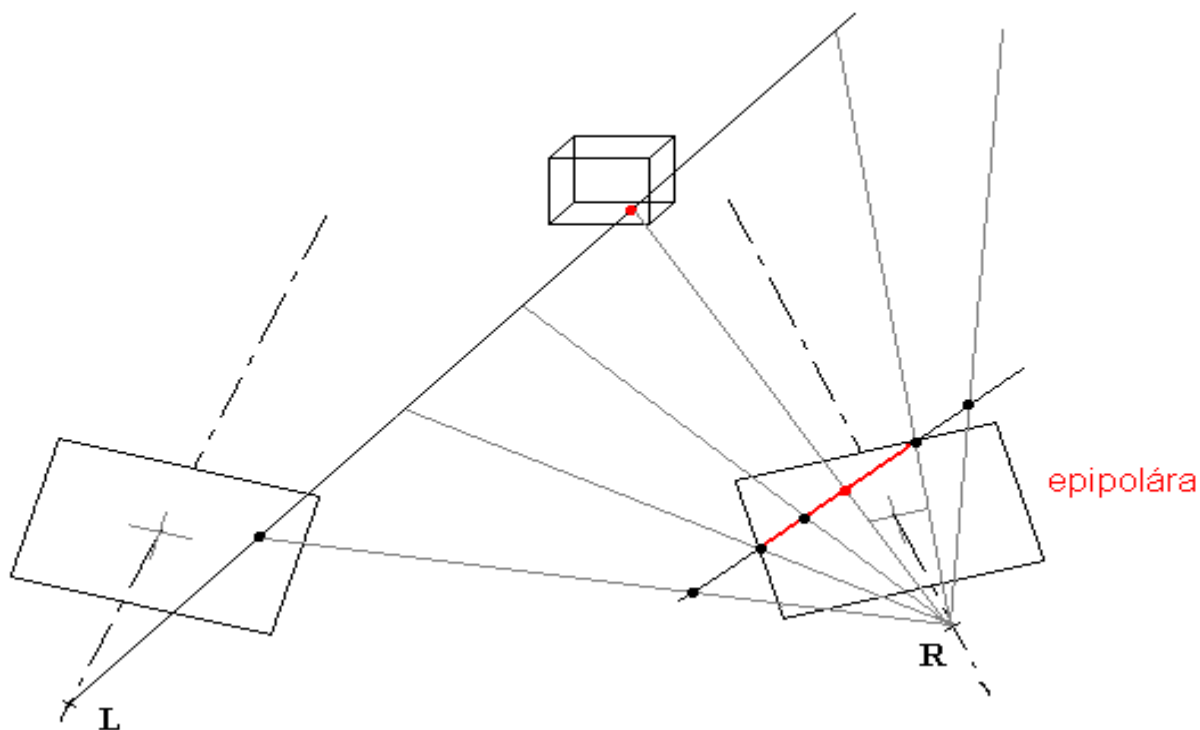


$$R = \begin{bmatrix} \cos\varphi \cdot \cos\kappa & -\cos\varphi \cdot \sin\kappa & \sin\varphi & 0 \\ \sin\omega \cdot \sin\varphi \cdot \cos\kappa + \cos\omega \cdot \sin\kappa & -\sin\omega \cdot \sin\varphi \cdot \sin\kappa + \cos\omega \cdot \cos\kappa & -\sin\omega \cdot \cos\varphi & 0 \\ -\cos\omega \cdot \sin\varphi \cdot \cos\kappa + \sin\omega \cdot \sin\kappa & \cos\omega \cdot \sin\varphi \cdot \sin\kappa + \sin\omega \cdot \cos\kappa & \cos\omega \cdot \cos\varphi & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Epipolární linie

- známe-li vzájemnou relaci mezi dvěma snímky, můžeme určit v jednom snímku polopřímku (u standardních kamer, jinak obecně křivku), která odpovídá bodu na druhém snímku
- této křivce se říká epipolární linie. Po určení epipolární linie nemusíme odpovídající si body hledat v celém snímku, ale pouze v okolí epipolární linie
- přímky (hrany) ve směru báze nenesou informaci (3D detekce hran se zlepší použitím trojice kamer)

(Pozn.: u uvedeného snímku je třetí kamera patrně „hledáček“)

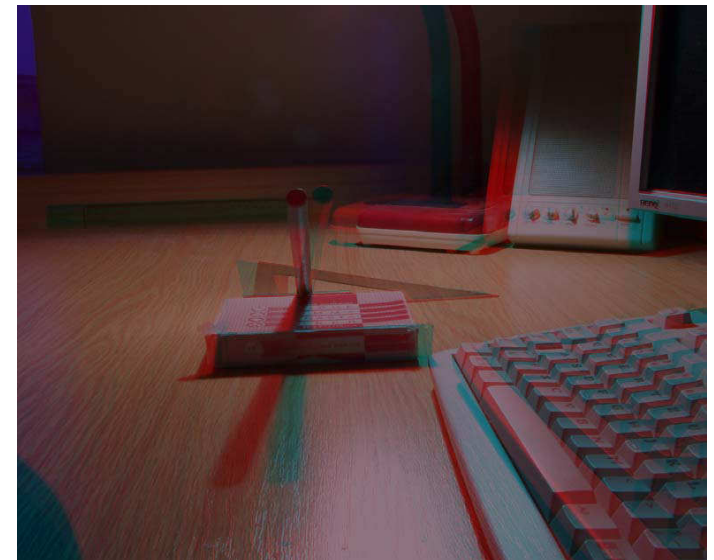
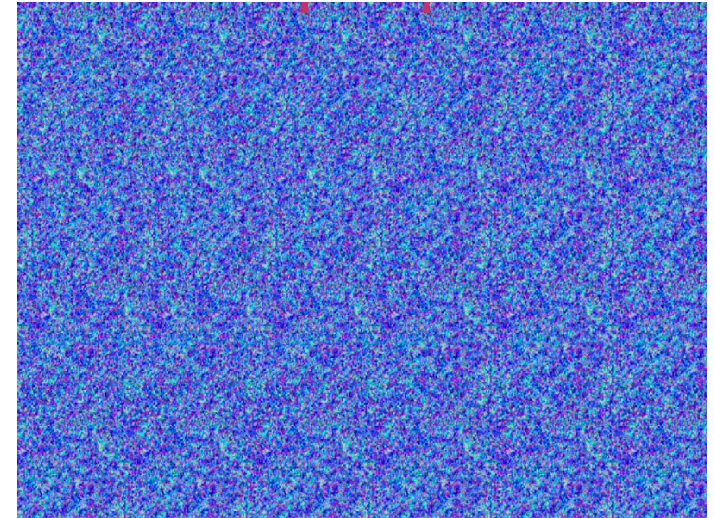


Rektifikace

- V současné době se často provádí zobrazení ve 3D
- 3D je možné zobrazit jako SIRDS (z hloubkové mapy) či jako anaglyph
- 3D zobrazení je také v 3D TV (aktivní, pasivní)
- pro tyto typy zobrazení je nutné snímky upravit tak, aby byly planoparalelní a měly stejné zvětšení. Jinak dochází k výraznému namáhání nervové soustavy při zpracování „posunutých“ dat
- pro korekci snímků se využívá tzv. rektifikace – snímky se převedou tak, aby ležely v jedné rovině a snímky měly vlastnosti jako by byly pořízeny kamerami se stejnými parametry a rovnoběžnými osami pohledu
- rektifikační vzorce pro levý a pravý snímek (P bod, M parametry původního zobrazení (kamera) a požadovaného průmětu, R transformace rektifikace a vzájemná mezi kamerami

$$P_{l,rect} = M_{new} \cdot R_{rect} \cdot M_l^{-1} \cdot P_l$$

$$P_{r,rect} = M_{new} \cdot R_{rect} \cdot R \cdot M_r^{-1} \cdot P_r$$



Pasivní 3D měření rychlosti vozidla

- orientační měření rychlosti projíždějícího vozidla
- v daném prostoru je již nainstalována jedna kamera. Bude na měření stačit?
- možná řešení – využít stávající kameru, přidat kameru druhou, nasvítit vzorem
- nasvětlení není možné z prostorových důvodů (přívod elektřiny, vzájemná poloha vůči kameře)
- zhodnocení možností měření jednou nebo dvěma kamerami



Rozbor přesnosti měření polohy pro objekt:

- pohybující se v dané rovině, jedna kamera
- pohybující se v okolí dané roviny, jedna kamera
- dvě kamery



v [km/h]	t [s]	Σ [m/s]	σ [km/h]
50	0,22	0,12	0,44
100	0,11	0,25	0,9
150	0,05	0,54	1,9

Standardní odchylka pro známou rovinu
pohybu bodu

v [km/h]	T [s]	σ [m/s]	σ [km/h]
50	0,22	1,36	4,9
100	0,11	27	9,8
150	0,05	6	21,6

Standardní odchylka pro známou rovinu a
toleranci $\pm 5\text{cm}$

v [km/h]	t [s]	σ [m/s]	σ[km/h]
50	0,22	0,5	1,8
100	0,11	1	3,6
150	0,05	2,2	7,92

Standardní odchylka určení rychlosti pro
stereofotogrammetrii – báze 0,3 m

Stanovení rychlosti vozidla

- pro zjištění polohy projíždějících vozidel byla zvolena metoda jedné kamery a plochy ve scéně
- ve scéně křižovatky bylo změřeno několik vzdáleností pro změřitkování
- byly pořízeny snímky křižovatky z různých směrů
- za pomoci kalibračního vzoru byly určeny parametry vnitřní orientace kamery pořizující snímky
- ze snímků byly zjištěny prostorové souřadnice bodů na povrchu vozovky
- ze změřených bodů na povrchu vozovky byla určena poloha měřící kamery a parametry její vnitřní orientace
- byla parametrizována rovina v dané vzdálenosti nad rovinou vozovky
- byla vypočtena transformace z roviny čipu (průmět bodu měřeného objektu pohybujícího se v dané rovině) na rovinu ve které se pohybuje bod
- po stanovení polohy měřeného bodu na čipu byla použita transformace, která ho převedla do soustavy vozovky (reálné souřadnice)



Zvolená varianta s jednou kamerou

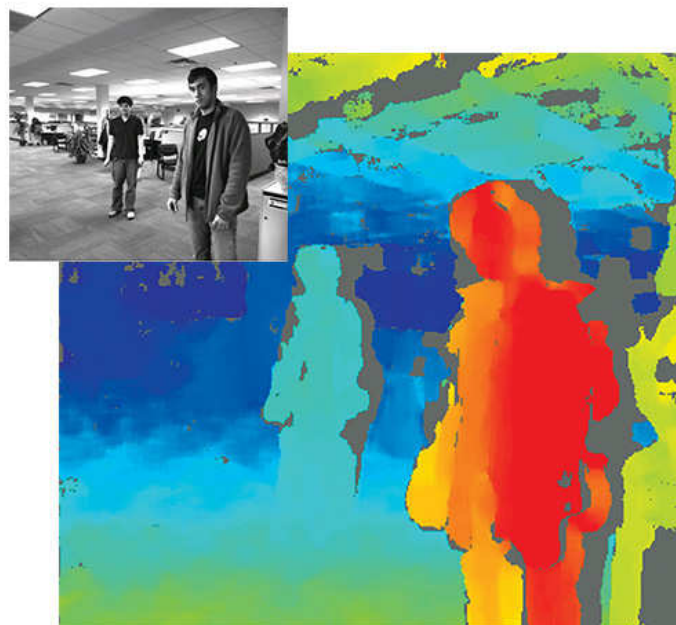
- metodu není možné použít, pokud neznáme, kde se pohybuje objekt
- jako „rovinu“ pohybu zvolíme výšku od vozovky, kde se pohybuje nárazník osobního vozidla
- lze-li výslednou plochu prohlásit za rovinu, je výsledkem transformace z roviny (čipu) na rovinu (pohybu) tělesa
- zvolenou rovinu je nutné změřit a parametrizovat pro následující výpočty. Je nutné znát její vzájemnou polohu s kamerou.
- výslednou polohu pohybujícího se objektu (auta) získáme jako průsečík paprsku procházejícího odpovídajícím/měřeným pixelem a danou rovinou
- postup stanovení měřící roviny: stanovíme rovnici reálné roviny země vůči kameře a tuto rovinu „zvedneme“ do požadované výšky, ve které budeme měřit.

Stanovení „roviny“ země

- díky provozu není možné provést přesná měření snímané scény (mechanickými měřeními)
- je možné provést pouze několik měření vzdáleností ve scéně
- ke stanovení parametrů scény je možné použít některou z metod počítačového vidění:
stereofotogrammetrie, obecná fotometrie

HW pro 3D počítačové vidění

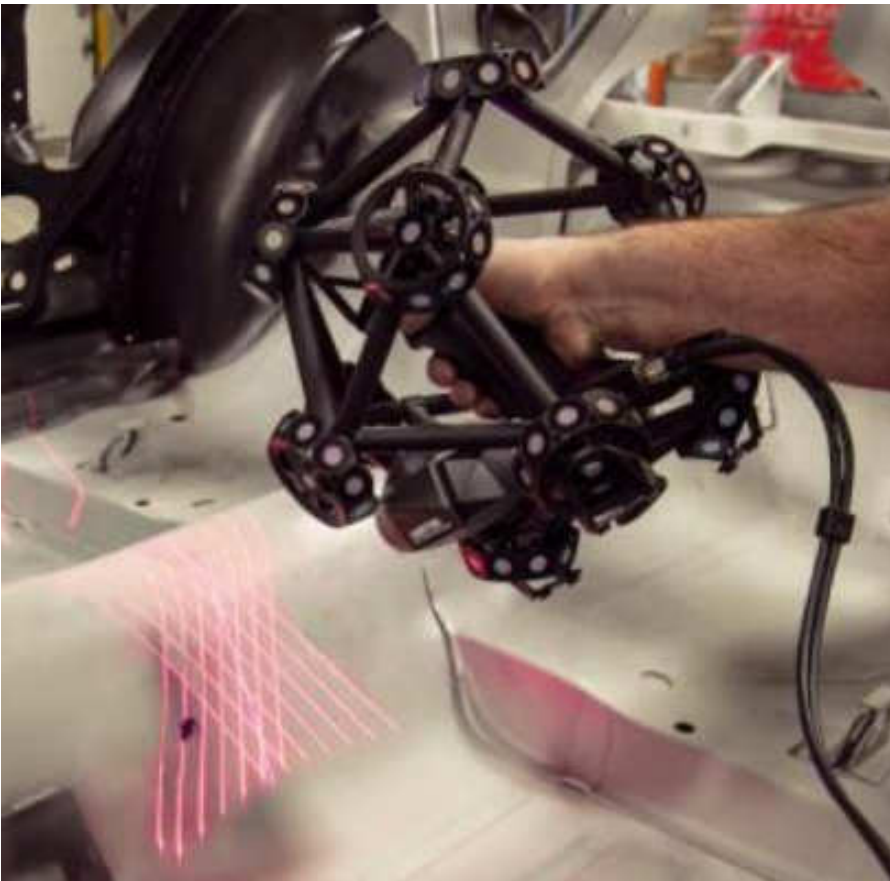
- stereo kamera bumblebee – Point grey (<http://ww2.ptgrey.com/stereo-vision/bumblebee-2>, <https://www.ptgrey.com/triclops>)



HW pro 3D počítačové vidění

- kombinovaný detektor vzdálenosti a 3D skener – (<http://www.creaform3d.com/en/metrology-solutions/optical-3d-scanner-metrascan> , <http://www.creaform3d.com/en/metrology-solutions/portable-3d-scanners> , <http://www.ndigital.com/products.php>)

0.2 – 10m; přesnost 0.03mm



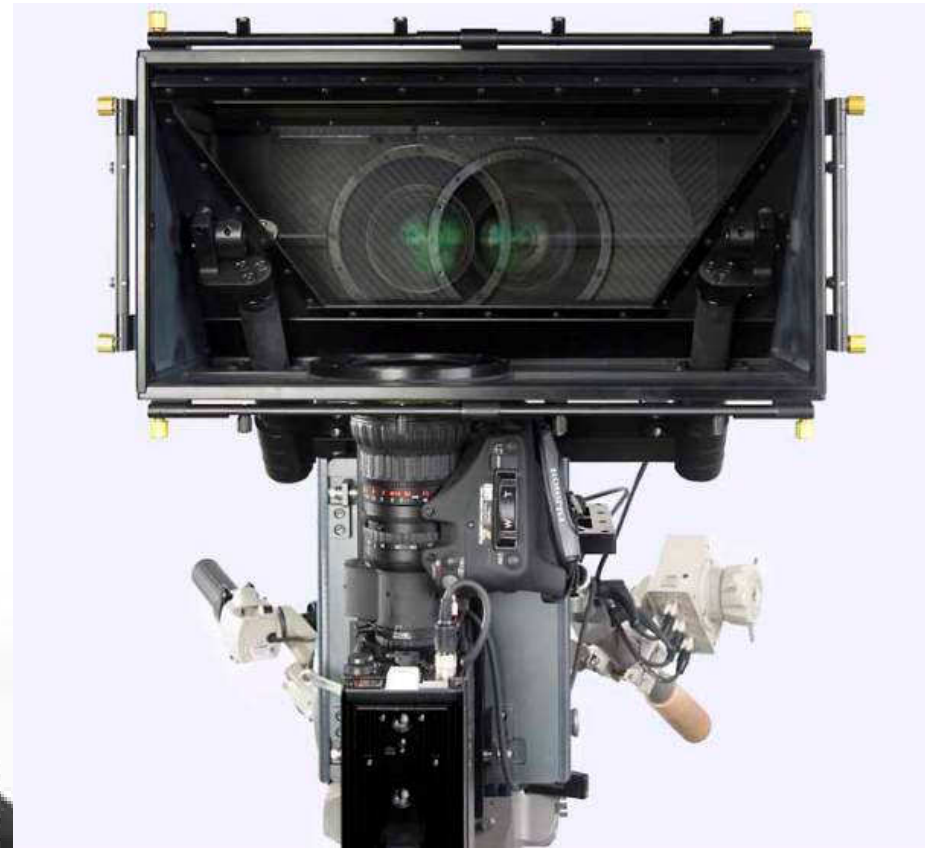
HW pro 3D vidění

- rozšíření (předsádky) běžných kamer o 3D možnosti, nebo přímo 3D kamery. Buď ukládají stereo snímky na jeden snímek (side-side), nebo přepínají Levý-Pravý. (Panasonic, Sony, Kúla)
- Fujifilm fotoaparát pořizuje stereo fotografie i videa
- nevýhodou je malá báze, omezující vzdálenost, při které je možno dosáhnout prostorového vjemu



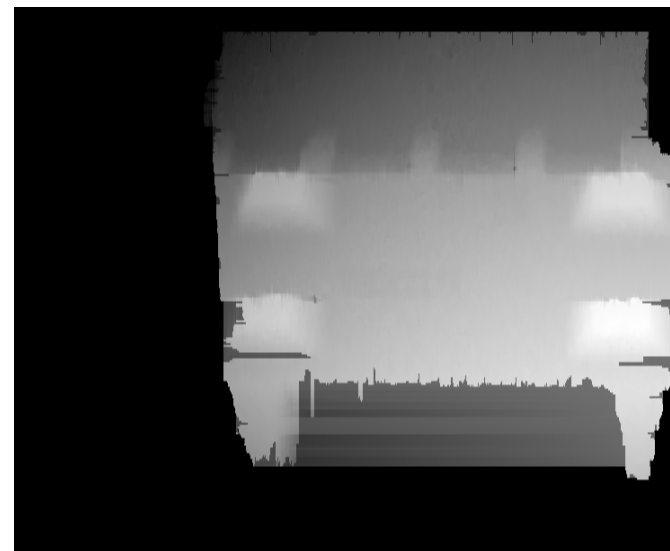
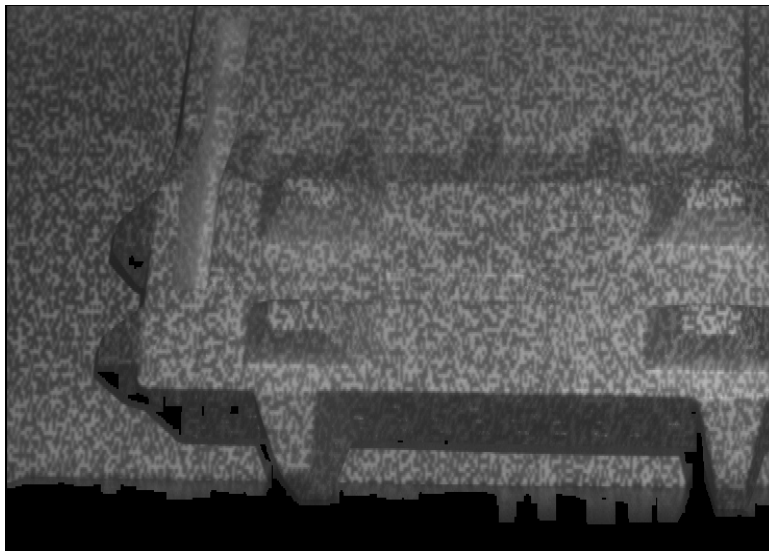
HW pro 3D počítačové vidění

- pro 3D TV – stereo kamery, kamerové soustavy (pole)
- cameron pace group (3D filmy) -
<http://www.cameronpace.com/v2/index.php/whatwedo/equipment/251>



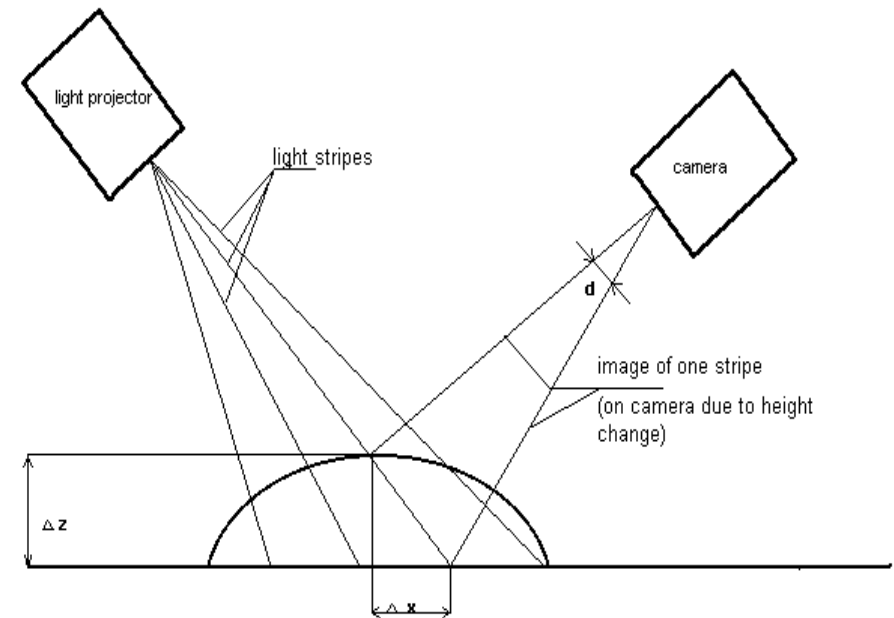
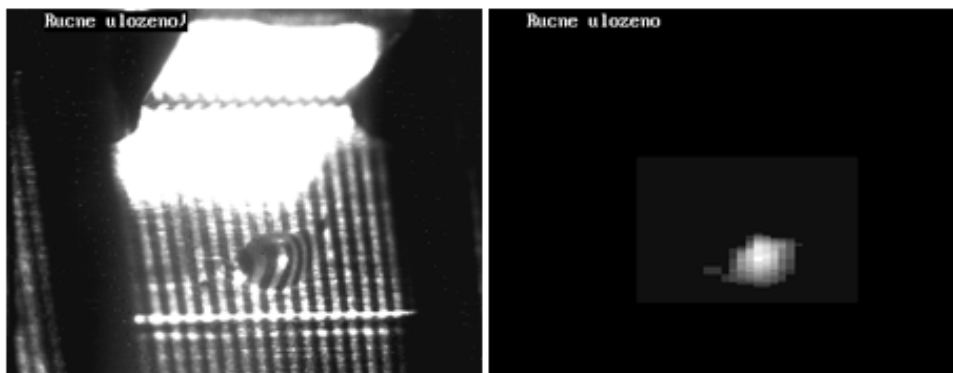
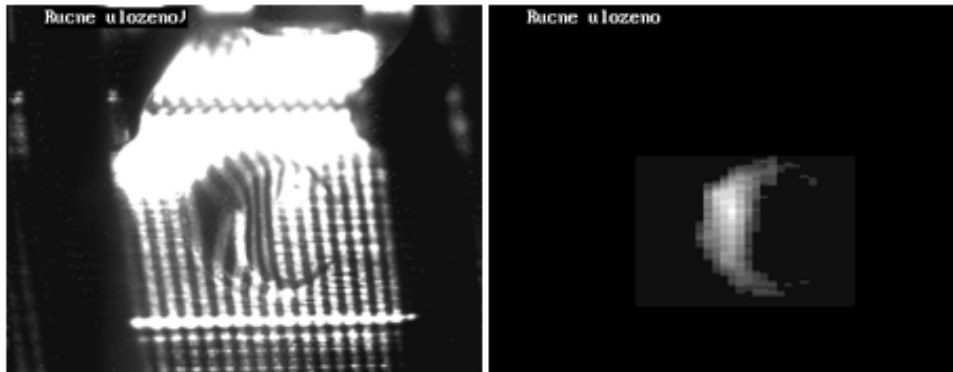
Aktivní metody 3D

- nasvětlení vzorem pro dodání bodů (hran) pro zvýšení počtu bodů pro vyhodnocení více kamerami
- používají se různé vzory – barevný vzor, černobílé šachovnice, pruhy různých roztečí (u statických scén), vzor s kódem (na základě střídání černobílých plošek v okolí je možné určit polohu ve vzoru – možnost lepšího sesouhlasení), nasvícení „bílým šumem“ (s úpravou větších ploch) – vlastně vzor, který by se neměl opakovat
- objekt nasvícen bílým šumem a snímán dvojicí kamer
- na základě nasvícených bodů je možné najít společné prvky i na rovinných úsecích, což by nebylo bez nasvícení možné
- výsledkem je hloubková mapa – nejsvětlejší část je nejvyšší (nejblíže kameře)

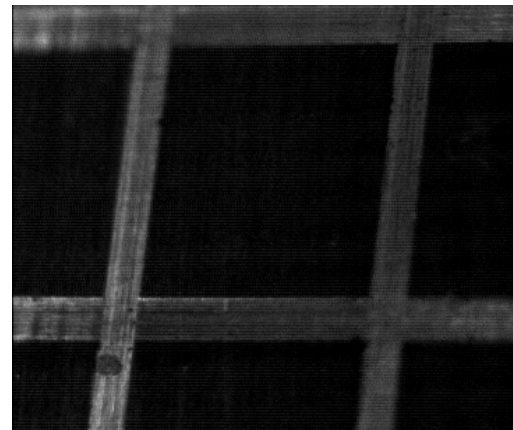
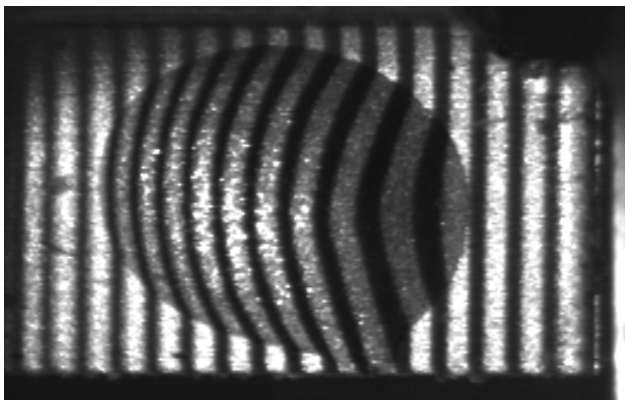


Aktivní nasvětlení vzorem jako součást měření

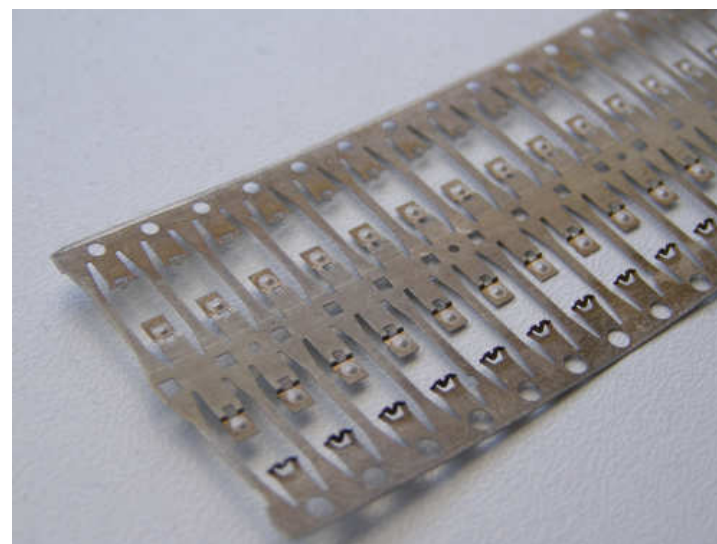
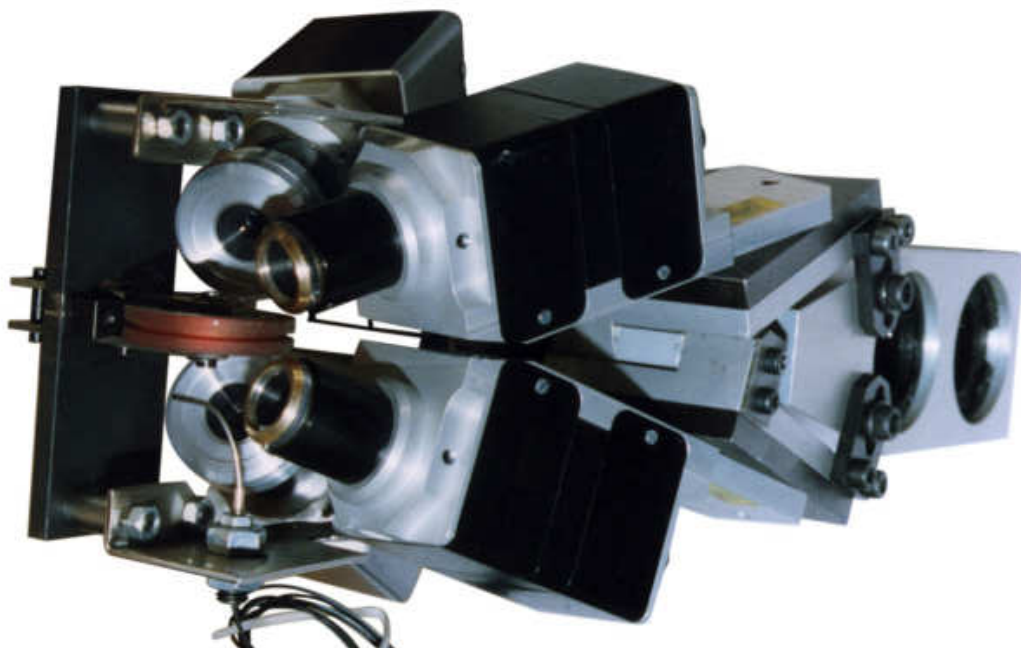
- 3D objekt (kapka) je nasvětlen soustavou proužků
- světlo je součástí měření (inverzní paprsky jako u detektoru – princip stejný)
- proužky se snadno detekují, vychýlení proužku nese informaci o výšce objektu



- nutno kalibrovat ve 3D – kalibr musí umožnit kalibraci ve výšce i ploše



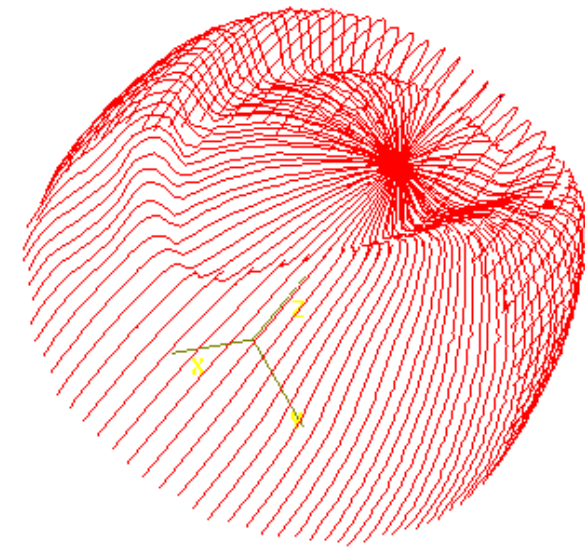
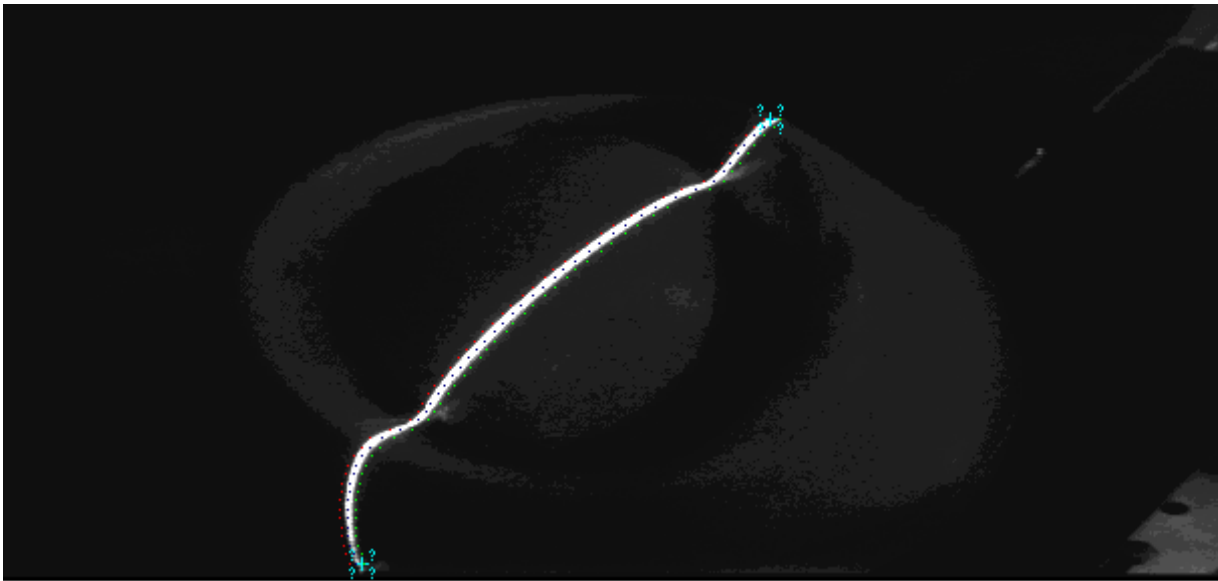
výsledný snímek nasvětleného objektu, snímek kalibru pro určení plošných rozměrů (bez nasvícení)



měřicí sestava (dva systémy) a pásek s měřenými součástkami (2x 45 součástek – měření do 5s)

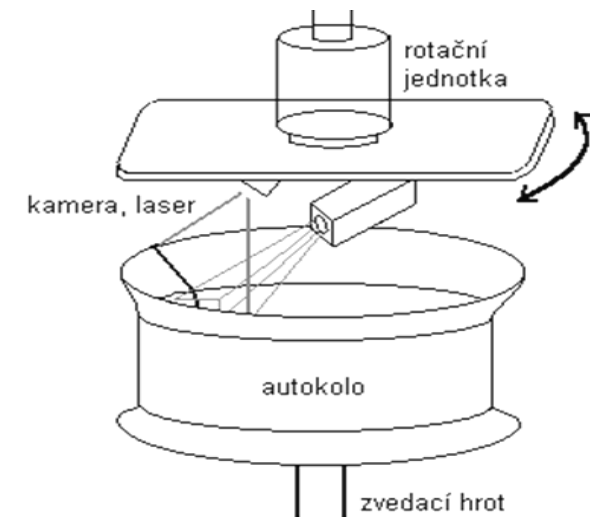
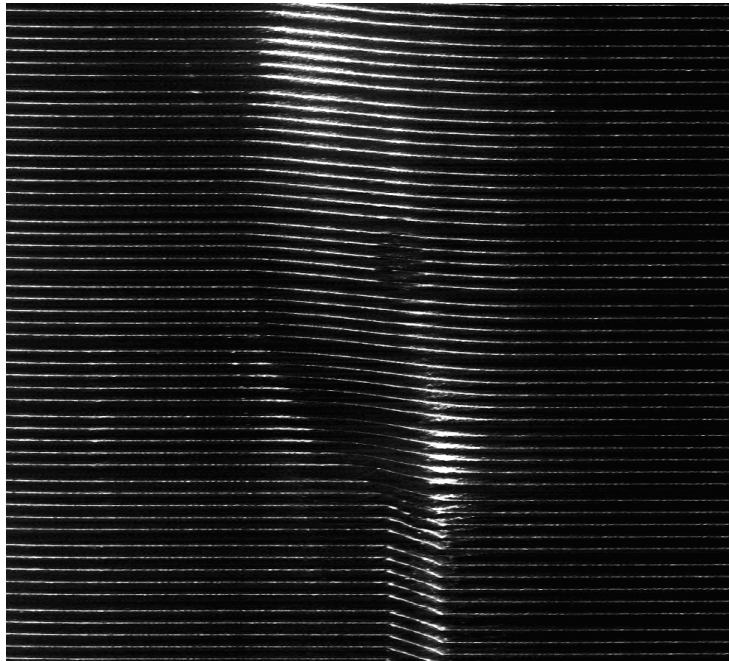
Aktivní měření pomocí laserové plochy

- nasvícení laserem dává dobré kontrastní poměry pro detekci. Laser je jasný. Pomocí filtrů lze odstranit spektrum patřící k pozadí
- metoda je pomalejší (na snímku je omezené množství světelných řezů)
- nutnost kalibrace světelného zdroje (například rovina rovina)
- nutná znalost pohybu (objektu nebo součástí měřícího systému)
- generátory laserového paprsku se prodávají v různé kvalitě (šířka svazku)



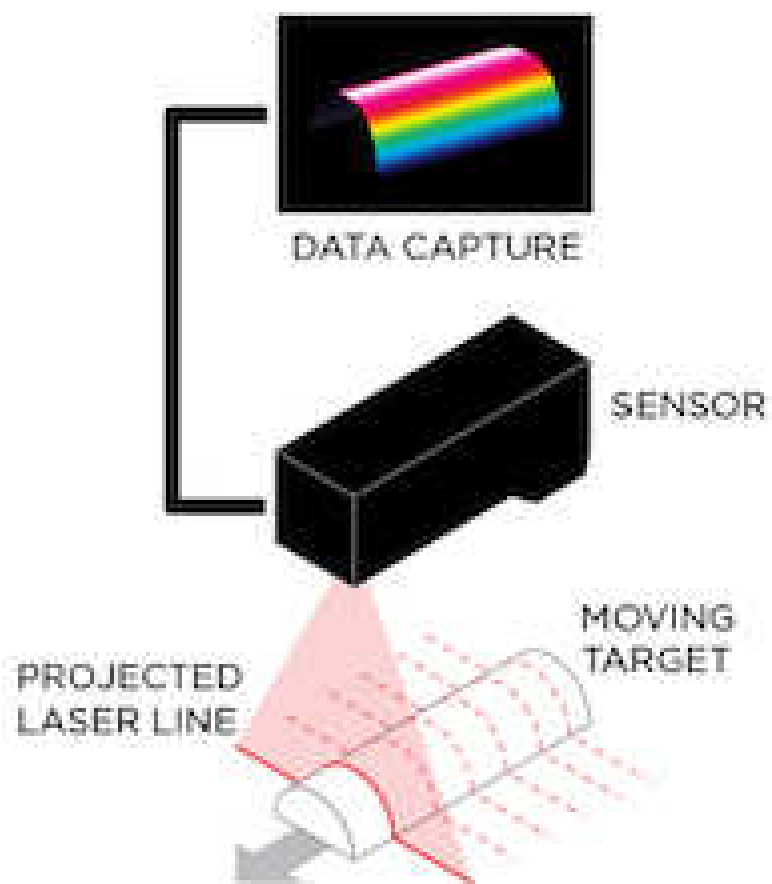
Měření defektů a rozměrů svarů

- pro zvýšení rychlosti měření jsou na měřící pozici čtyři systémy
- využívá se plošné kamery s vícenásobnou expozicí - na jednu expozici se vyčte pouze několik řádků obrazu
- detekují se chyby ve svaru (kapky, bubliny, svařovací drát ...) a velikost (délka, šířka a výška svaru)



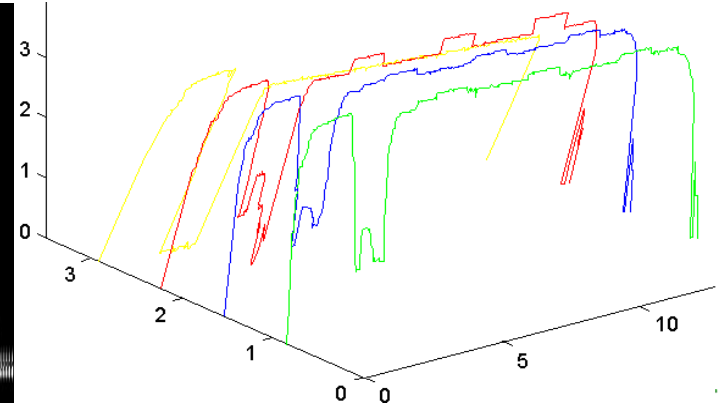
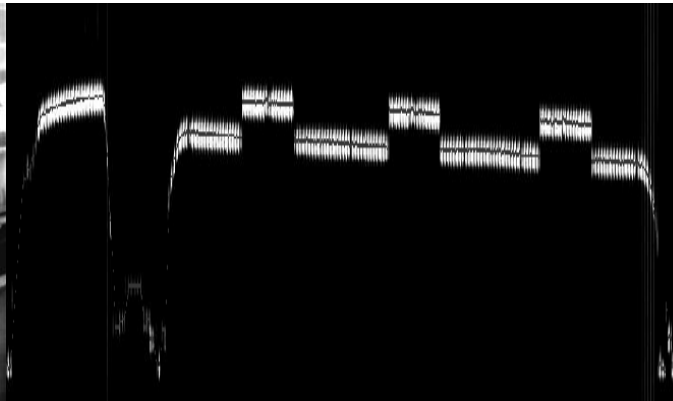
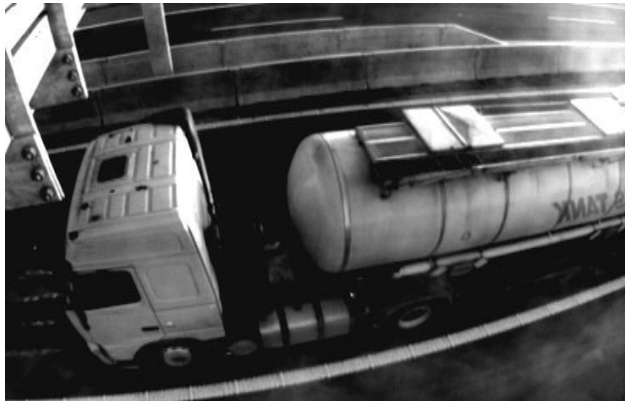
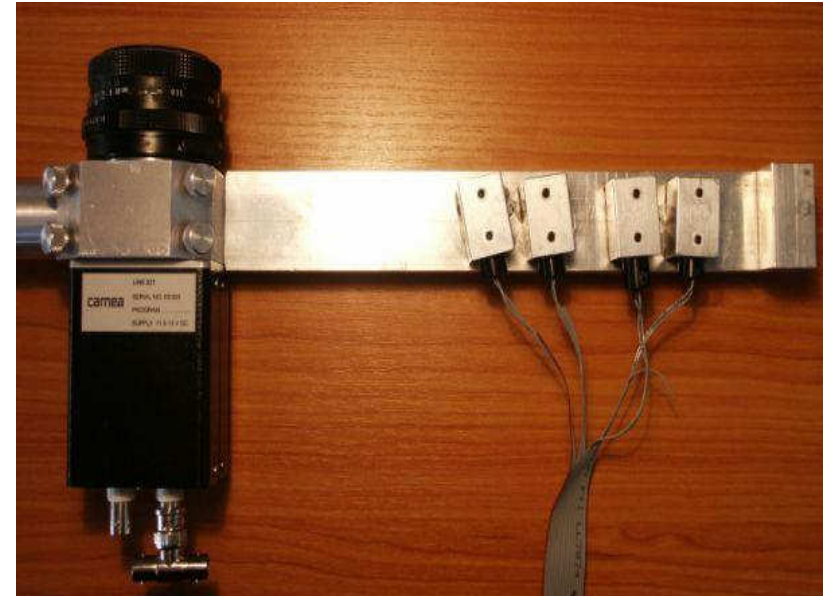
HW pro 3D počítačové vidění

- laserový pruh gocator - (<http://www.lmi3d.com/products/gocator/>)



Měření kategorií automobilů na základě profilu

- měří se pomocí lineární kamery (expozice/snímek v řádu us)
- scéna nasvícena několika laserovými paprsky
- na základě paprsků se získá profil, který se vyhodnotí a určí se třída automobilu
- za pomoci dvou systémů posunutých ve směru jízdy, lze určit i přibližnou rychlost vozidla
- systém se kalibruje v laboratoři



HW pro 3D počítačové vidění

- Microsoft Kinect <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.405.8789&rep=rep1&type=pdf>
<https://www.researchgate.net/publication/308615218> Comparison of Kinect v1 and v2 Depth Images in Terms of Accuracy and Precision

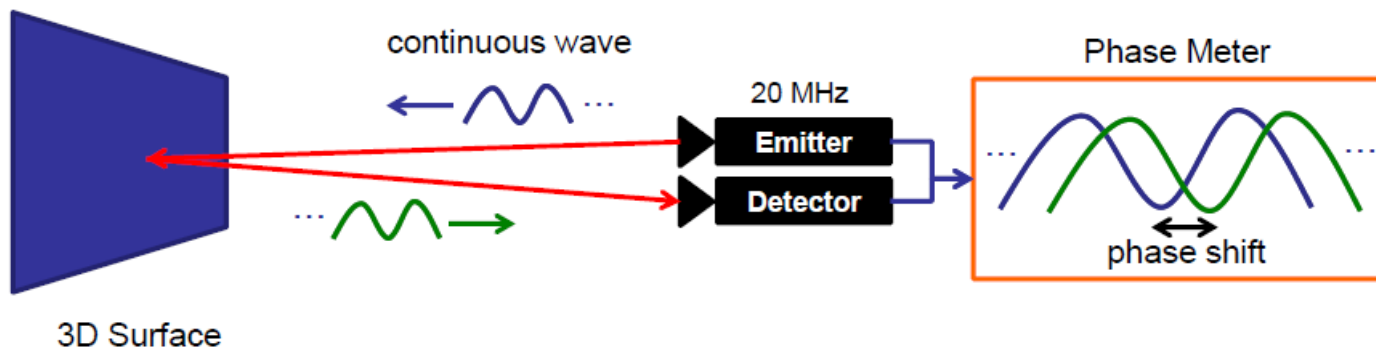
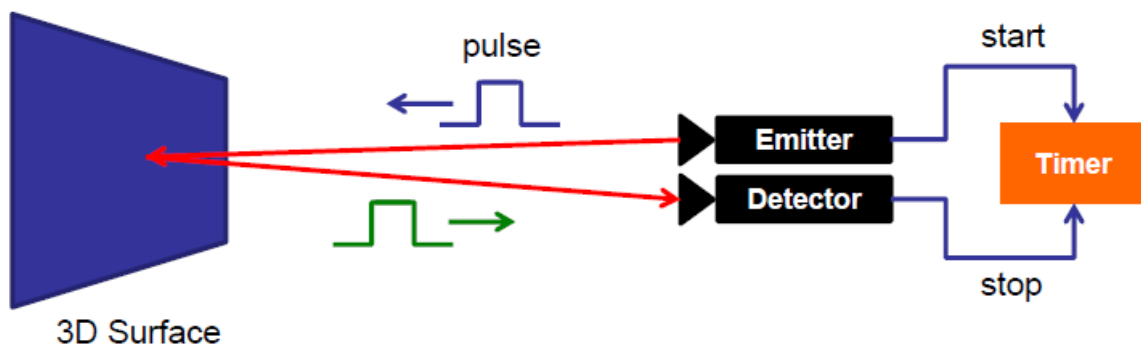
- Verze 1

IR nasvícení vzorem, vzorem jsou body (speckles), body jsou náhodně rozmístěny, „shluky“ bodů jsou jedinečné (bílý šum), tvar bodů se mění se vzdáleností, měření rozděleno do tří zón, Zpracování a určení výsledného tvaru je náročné, 30 Hz, 640x480, nutná kalibrace kamery a projektoru,



- Verze 2

TOF (Time Of Flight), jasové i vzdálenostní/hloubkové informace společně, rychlejší zjištění vzdálenosti, méně přesné (nutné přesné určení času, náběžné hrany nejsou ostré), 30Hz, 512 x 424



HW pro 3D počítačové vidění

- pmd 3D senzor – TOF - ifm (<https://www.ifm.com/gb/en/shared/technologies/3d-smart-sensor-o3m/applikationen/abstandskontrolle>)

Interferenční metody

- detekce nerovností na plošných (spojitých) objektech
- detekuje nerovnosti v řádech délky vlny použitého záření (změna vzdálenosti o délku vlny = vzdálenost od maxima jasů k sousednímu maximu)
- detekce rovinných deformací na silikonové destičce
- https://www.researchgate.net/publication/298835725_Optical_methods_in_experimental_mechanics



HW pro 3D počítačové vidění

- lightfield - <http://graphics.stanford.edu/projects/lightfield/>
Lytro <https://www.lytro.com/>
[Ren Ng, Digital Light Field Photography. Stanford PhD Dissertation. June 2006.](#)
- díky složitějšímu záznamu (více informace díky rozložení průmětny pomocí microlens) je schopen provést rozsáhlejší zpracování (přeostřovat)

