

# **Západočeská univerzita v Plzni**

**Fakulta aplikovaných věd**

**Katedra Kybernetiky**

## **3D skener na bázi strukturovaného světla**



**Student: Jiří Hlaváček**

**Vedoucí práce: Ing. Zdeněk Krňoul, Ph.D.**

**Datum: 20.8.2013**

## **Abstract**

*Reconstruction with the help of a camera and structured light is a modern approach to the problematics of Computer Vision and development in this field goes rapidly up. If we have a calibrated camera-projector system, where the source of structured light projects onto our examined object and we possess images containing these projections, it is possible to find correspondence points between every single image. From these correspondence points it is then possible with the triangulation method to determine, where these points in the image plane lie in world coordinates. We analyze in more detail imaging with Binary patterns, for greater understanding of the problems of subject for future application. Our goal is to find a ideal system and its description, which would be cheap to construct and easy to relocate.*

## **Abstrakt**

*Rekonstrukce pomocí kamery a strukturovaného světla je moderní přístup k problematice počítačového vidění a vývoj v tomto směru jde rapidně nahoru. Máme-li kalibrovaný systém kamera-projektor, kde zdroj strukturovaného světla promítá na zkoumané těleso a máme snímky této projekce, je možné najít korespondenční body mezi jednotlivými snímky. Z těchto korespondenčních bodů je pak možné metodou triangulace zjistit, kde body ze snímku v ploše leží v prostoru. Analyzujeme podrobněji snímání pomocí binárních vzorů, pro nastudování problematiky této disciplíny pro budoucí aplikaci. Cílem je nalézt vhodný systém a jeho popis, který by byl levný na sestavení a snadný na přenos.*

# Obsah

## Abstrakt

### 1 Úvod

1.1 Obecný postup

1.2 Podmínky

### 2 Systémy 3D skenerů

2.1 Systém využívající dvě kamery – Stereovision

2.2 Systém kamera – projektor

2.2.1 Vzory

2.2.2 Binární vzor

2.2.3 Realizace

### 3 Kalibrace

3.1 Kalibrace kamery

3.1.1 Zhangova kalibrační metoda

3.2 Kalibrace projektoru

3.3 Výsledky kalibrace

### 4 Shrnutí

### 5 Ukázky a přílohy

## Literatura

# 1 Úvod

Cílem této bakalářské práce je seznámení se problematikou počítačového vidění a pokus o sestavení jednoduchého 3D skeneru za pomoci jediné kamery a projektoru. Tento skener má výhodu snadného přenosu a dostupnosti, proti fixním skenerům. Dnes existuje několik metod pro získání prostorové informace o objektech v reálném světě, ale každá metoda je vhodná pro různé problémy. Problematika je určena podle toho zda se snímaná scéna smí pohybovat či ne, nebo zda chceme rekonstruovaný objekt zachovat i s informací textury. Zaměřili jsme se na nepohyblivé objekty a jejich zpracování pomocí binárního vzoru pro získání potřebných znalostí v další práci. Binární metoda je již dobře známý postup, který je dobře zdokumentovaný a pro začátek studia problematiky nám poslouží dobře. Ostatní metody využívající jiné vzory analyzoval a popsal v [1] Salvi se svým studentem Navarrezem. Pro naše záměry jsme zvolili bystu jako náš nepohyblivý zkoumaný objekt. Tato volba byla učiněna, abychom viděli problematiku s rekonstrukcí lidského obličeje. Bysta ale od člověka má tu výhodu pro náš výzkum, že se nemůže pohybovat a materiál má také jiné vlastnosti než lidská kůže. Předpokládáme, že by bylo možné tímto postupem snímat lidské obličeje a v budoucnu se na tuto téma zaměříme v další práci.

## 1.1 Algoritmus

Obecný postup pro postup v počítačovém vidění můžeme rozdělit do několika kroků. Pokud je metoda pro skenování zvolena je možné přejít ke kroku kalibrace, která je samotným počátkem procesu. Po úspěšné kalibraci přecházíme ke snímání objektu danou metodou. Snímání a kalibrace jsou kroky nejnáchylnější k chybám vlivem lidského faktoru. Dalším krokem je tedy analýza snímků a jejich případná rektifikace, případně jejich vyloučení ze souboru dat. Z tohoto souboru dat provádíme

následně identifikaci každého pixelu ve snímku a přiřadíme mu jednoznačně označení informace. Zpracováním bodů jsme dokončili práci se snímky. Přichází krok rekonstrukce z informací v jednotlivých souřadnicích. Po rekonstrukci máme 3D mapu zkoumaných bodů v prostoru a můžeme porovnávat výsledky s výsledky jiného skeneru, nebo jednoduchým pohledem na výsledné zrekonstruované těleso a reálný objekt. Přesnější evaluace výsledku je pomocí již nasnímané 3D mapy objektu, ale ne vždy je toto možné.

## **1.2 Podmínky**

Během celého procesu se nesmí jakkoliv manipulovat s vlastnostmi užitých zařízení, aby bylo možné výpočtově řešit problematiku. Proměnlivost parametrů znemožňuje výsledky a vztahy zavedené v této práci. Všechny metody dnes známé a používané vychází z této jednoduché podmínky. Je potřeba mít základní znalosti lineární algebry. Snímky s vyšším rozlišením obsahují mnoho bodů a je potřeba uvážit i tento faktor. Snímky s vysokým rozlišením sice mohou poskytovat hodně přesné informace při hledání korespondentů a zkoumání rozdílů intenzit, ale ne vždy je toto prospěšné. Metody rozčlení snímky na určité segmenty a přiřadí celému segmentu kód. Nepřináší tedy další informace, ale prodlužují výpočtový čas celého procesu. Správná volba rozlišení je důležitá apriorní úvaha metod.

## 2 Systémy 3D skenerů

Systémy využívané pro rekonstrukci 3D mapy povrchu objektu dělíme na kontaktní a nekontaktní. Kontaktní systémy (contact techniques) projíždí kalibrovaným senzorem každý bod povrchu tělesa a měří odchylky. Senzor musí být v kontaktu s každým bodem a to způsobuje velikou časovou náročnost. Dále jsou tyto systémy drahé a také jsou náchylné k opotřebování pohybem senzoru. Senzor se musí po každém měření kalibrovat. Tyto systémy jsou první, které byly využívány pro rekonstrukci objektů v počítačovém vidění. Nekontaktní systémy (non contact techniques) začaly být vyvíjeny, aby odstranily nedostatky kontaktních systémů. Senzor je nahrazen jednou či více kamerami. Tento přístup zkoumání objektu lze dále rozdělit na pasivní a aktivní. Pasivní postup snímá scénu z různých úhlů a pak hledá korespondenty mezi snímky. Je možné takto využít jednu nebo více kamer, ale je nutné mít použítá zařízení před snímáním kalibrovaná a znát vztah mezi jednotlivými kamerami nebo pohyb kamery pokud používáme pouze jednu. Tato metoda rekonstrukce je závislá na schopnosti nalézt korespondenty mezi jednotlivými snímky, které nemusí být vždy snadné k nalezení. Aktivní přístup nahrazuje druhou kameru aktivním zařízením (projektor), které promítá vzor strukturovaného světla na objekt. Promítnutí vzoru (pattern) na objekt jednoznačně určuje korespondenty a dává jednotlivým bodům kód nutný pro rekonstrukci 3D mapy povrchu objektu. Apriorní znalost promítnutého vzoru dovoluje rekonstrukci pomocí jediného snímku v některých postupech, porovnáním vzoru a jeho deformací na snímku objektu. Pro systémy využívající kameru a projektor je několik různých vzorů a jejich variace. Jednotlivé vzory mají své specifické využití podle potřeby, daného problému či objektu a dostupného zařízení. Dále se hlouběji podíváme na systém dvou kamer pro lepší pochopení postupu a jeho problematiky a systému s kamerou a projektorem, který je hlavní náplní této práce.

## 2.1 Systém využívající dvě kamery – Stereovision

Stereovision je pasivní přístup pro nekontaktní systém, který funguje stejně jako lidský zrak. Dvě lidské oči jsou zde reprezentovány soustavou dvou kamer, jejichž vztah je předem znám a parametry kamer jsou určeny jejich kalibrací na objektu, jehož vlastnosti jsou apriorně známé. Aby byla možná rekonstrukce, tak obě kamery nesmí být v žádném případě vůči sobě rovnoběžné. Tato podmínka je esenciální pro rekonstrukci metodou triangulace. Z jednoho snímku je možno určit dvě souřadnice pro jeden bod, ale k tomu abychom určili třetí prostorovou souřadnici potřebujeme druhý snímek. Z prvního pohledu vidíme na jaké přímce leží daný bod v prostoru a druhá kamera vidí tuto přímku z boku a nalezne daný bod. Takto se sestaví trojúhelník, ze kterého pak vypočteme souřadnice v prostoru jednoho bodu. Je tedy důležité správně nalézt jeden konkrétní bod v obou pohledech. Tyto body shodující se v obou snímcích se nazývají korespondenty a jejich nalezení je hlavní náplní této metody. Bez druhého pohledu není možné transformovat scénu do prostorových souřadnic. U člověka je tento problém stejný. Lidé pouze s jedním okem mají potíže s prostorovým vnímáním.



*Fig 2.1: Ukázka stereovision sestavy*

Stereovision může pracovat i s jednou jedinou kamerou. Cena sestavy tedy by byla téměř poloviční, ale princip stereovision musí být zachován. Druhý pohled je zajištěn první kamerou, která se přesune na pozici druhé kamery. Tato pozice musí být dopředu

známa pro vytvoření vztahů transformace do 3D souřadnic. Na obrázku [2.2] vidíme popsané vztahy mezi souřadnicemi mezi levým a pravým snímkem. Z tohoto obrázku je také vidět jak z ohniska levé kamery  $O_L$  vede přímka bodem  $X_L$ . Na této přímce leží bod  $X$  v prostoru. Bod  $X_R$  je korespondent bodu  $X_L$  a stejně jako na snímku levé kamery vedeme přímku z ohniska pravé kamery a v ideálním případě nalezneme průsečík přímek, ve kterém leží náš bod  $X$  a získáme dost informací pro výpočet polohy v prostoru.

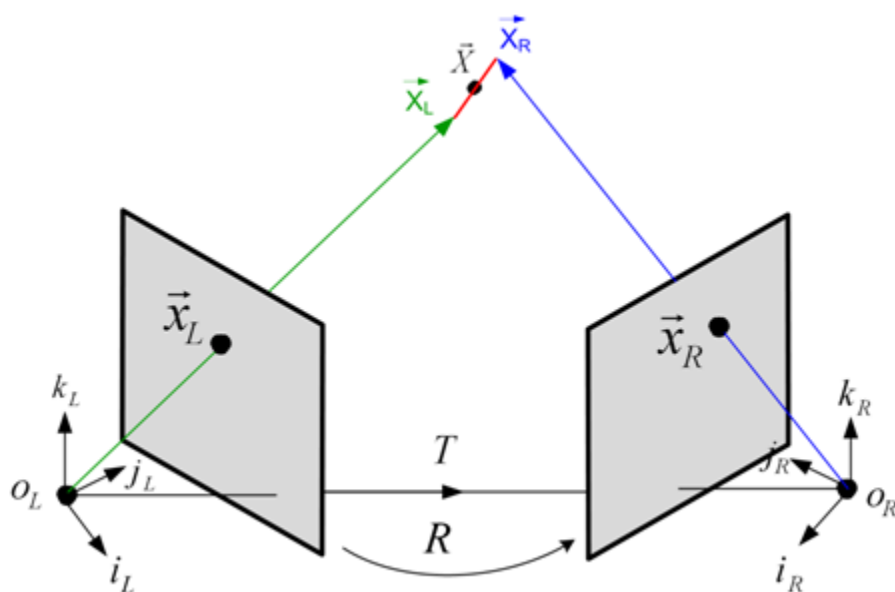


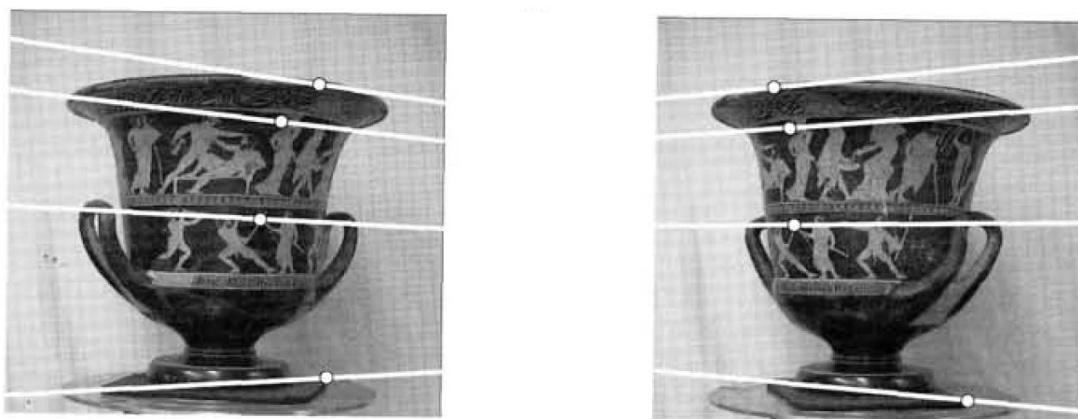
Fig. 2.2: Pohled na triangulaci a poloha kamer

Na obrázku vidíme, že nemusíme nalézt průsečík a je třeba provést mezi výpočet minimální vzdálenosti přímek. Bod  $X$  pak určíme ve středu úsečky, která je minimální vzdálenost přímek. Vektory  $T$  a  $R$  jsou vektory translace (posun) a rotace (natočení) popisující vztah mezi levou a pravou soustavou souřadnic kamer. Vztah mezi souřadnicovými systémy obou kamer je popsán fundamentální maticí. Fundamentální matice popisuje algebraicky vztah mezi korespondenty na obou snímcích. Musí platit vztah

$$x'Fx = 0 \quad (1)$$



kde  $x$  a  $x'$  jsou souřadnice korespondujících bodů mezi snímky s homogenními souřadnicemi. Díky tomuto vztahu pak lze určovat epipolární plochy [2.3]. Tyto plochy se vyznačují tím, že nám na snímcích ukáží přímky, na kterých leží korespondující bod z jednoho snímku na druhém. Celý princip stereovision tedy je závislý na nalezení korespondentů. V prostředí MatLab jsou funkce pro nalezení korespondentů (`matchFeatures`), pro výpočet fundamentální matice (`estimateFundamentalMatrix`) a pro znázornění epipolárních přímků ve snímku (`epipolarLine`).



*Fig. 2.3: Průsečky epipolárních ploch a roviny snímku*

## 2.2 Systém kamera – projektor

V této práci jsme se zaměřili na systém s jedinou kamerou a pro získání prostorové informace slouží projektor [2.4], který na těleso promítá určitý vzor. Rekonstrukce může pracovat s apriorní informací promítnutého vzoru a jeho deformací nebo pomocí kódu, který přiřadíme každému pixelu. Tyto metody se dále rozlišují na jednosnímkové a vícesnímkové. My jsme se rozhodli pro vícesnímkovou metodu pomocí binárních vzorů. Vícesnímkové metody nemohou snímat pohyblivé objekty, ale poskytují vysokou hustotu zrekonstruované 3D mapy při vyšším počtu snímků. Při

malém počtu snímků jsou nejvíce náchylné k chybám. Podíváme se na různé vzory strukturovaného světla používané pro tuto disciplínu.

### 2.2.1 Vzory

Strukturované světlo lze rozdělit na různé vzory. V tabulce 2.1 vidíme rozdělení vzorů na diskrétní (discrete) a spojité (continuous). Diskrétní vzory využívají jeden vzor nebo soustavu profilů, které se postupně promítají na objekt a každý vzor je snímán zvlášť. Spojité využívají funkčního popisu vzoru a spojitě mohou měnit parametr. V těchto metodách je využívána sinusová funkce a parametrem měnícím se v čase je fáze. Dále dělíme vzory podle způsobu kódování bodů. Podle zvolené metody tvoříme kódové označení každého bodu (codeword). Toto kódové označení je posloupnost hodnot, které nám slouží k výpočtu prostorové souřadnice bodu. Způsob získání tohoto kódu dělí vzory na prostorové a časové multiplexování. Prostorové generování kódu využívá okolních bodů zkoumaného pixelu a podle intenzity, textury nebo jiných způsobů filtrů. Časové multiplexování využívá informace pouze jednoho jediného bodu. Je potřeba mít soustavu snímků na kterých je vzor pozměněn a posloupnost kódu určí intenzita daného bodu v každém snímku. Prostorové multiplexování tedy vyžaduje jen jeden snímek, kdežto časové multiplexování vyžaduje minimálně dva snímky a aby bylo dosaženo větší hustoty bodů 3D mapy je počet použitých snímků vyšší a tím je znemožněno snímání pohybujících se objektů. Frekvenční multiplexování shrnuje všechny metody, kde fázové dekódování je prováděno frekvenčně a nikoliv prostorově. Pro tento pokus jsme se zaměřili na binární kódování.



V tabulce 2.1 vidíme souhrn několika vzorových metod používaných ve strukturovaném světle. V prvním sloupci vidíme, zda je metoda diskrétní či spojitá. Ve druhém sloupci máme informaci o tom kterou metodou multiplexování je určen kód bodu ve snímku. Třetí sloupec poskytuje náhled na vzory. Vedle každého vzoru je pak seznam autorů, kteří ve svém výzkumu daný vzor použili a ve kterém roce zveřejnili svou metodu rekonstrukce s jeho pomocí. U každého autora je také uveden ve sloupci *shots* minimální počet snímků potřebných k aplikaci postupu a ve sloupci *camera* je počet kamer použitých. *Axis* nám říkají v kolika osách se projevuje daný vzor. Je však zřejmé, že pouze některé prostorové metody multiplexování vyžadují druhou dimenzi a v tabulce je to potvrzeno. *Pixel depth* udává informaci o zpracování pixelů v barevných spektrech. *B* (black) znamená zpracovávání pouze v černo-bílém spektru. *G* (gray) zpracovává body pomocí odstínů šedi. *C* (color) může pracovat i s barvami. *Coding strategy* je způsob generování kódu pixelu. Generovat lze periodicky posloupností snímků v čase (periodic) nebo sadou snímků z více kamer z jednoho časového okamžiku (absolute).

### 2.2.2 Binární vzor

Metoda binárních vzorů (12inary patterns) postupně promítá na objekt černé a bílé pruhy a codeword každého bodu je určeno podle posloupnosti intenzity ve snímcích. Kódový řetězec pro bod pak může vypadat takto, když použijeme  $N$  snímků:

$$S(x, y) = \{B_1(x, y), B_2(x, y), \dots, B_N\} \quad (2)$$

Označení  $S$  je zvoleno jako zkratka pro slovo stripe (pruh) a hodnoty  $B_N$  nabývají hodnot  $\{0,1\}$ . Kód pixelu má tedy  $2^{n-1}$  bitů, kde  $n$  je počet

promítnutých vzorů na objekt. Začíná se vzorem obsahujícím jeden černý a jeden bílý pruh a každý další promítaný vzor má dvojnásobný počet pruhů. Důležitým aspektem této metody je mít co nejostřeji nastavený projektor. Je to z důvodu, že body na mezi černého a bílého pruhu se nejhůře přiřazují. Tento problém se jednoduše odstraní vyšším rozlišením snímků, ale je třeba mít projektor nastaven na totožné rozlišení. Metoda binárních vzorů na rozdíl od ostatních metod nepřiznává přesnou intenzitu bodu ve snímku, ale vlivem chyb a odrazů světla přiřazuje pouze hodnoty 1 a 0. Hodnotu 1 přiřazuje bodům ležícím v osvětleném bílém pruhu a hodnotu 0 dává bodům ležícím v tmavých pruzích. Pro každé těleso je ale důležité zavést vztah pro intenzitu z důvodu různých vlastností materiálů u různých objektů. Lidská kůže má jiné světlo absorpční vlastnosti než sádrová bysta nebo černá krychle. Je tedy nutné snímaný objekt nejprve plně osvětlit světlem stejné barvy a intenzity z projektoru jako je použito u bílých pruhů, abychom měli maximální intenzitu a pak těleso necháme neosvícené jako v případě tmavých pruhů pro minimální intenzitu. Rozdíl těchto intenzit je pak hodnota, kterou bychom měli dělit intenzity bodů ve snímcích. Dostáváme vztah:

$$J_K(x, y) = \frac{I_K(x, y) - I_L(x, y)}{I_H(x, y) - I_L(x, y)} \quad (3)$$

$J_K$  nám tedy dává hodnoty v rozmezí  $\langle 0, 1 \rangle$ .  $I_H$  je hodnota intenzity bodu  $X$  při maximálním osvětlení.  $I_L$  je hodnota bodu  $X$  při minimálním osvětlení. Tento vztah nám udává normalizovanou intenzitu snímku. V některých případech je možné  $I_L$  přiřadit hodnotu 0, ale ne vždy je toto vhodné. Tato možnost je závislá na osvětlení prostředí ve kterém je objekt snímán. Podle hodnoty  $J_K$  pak určíme, zda bod bude mít hodnotu ve snímku 1 nebo 0. Problém je určení bodů kolem hodnoty  $J_K$  rovno 0.5.

$$B_K(x, y) = \begin{cases} 1 : J_K(x, y) > 0.5 \\ 0 : J_K(x, y) \leq 0.5 \end{cases} \quad (4)$$

Kolem těchto bodů se nejčastěji projevují chyby a je třeba jako v každém reálném systému přidat chybu  $\xi$ . Zavádí se tedy vztah širší se započteným šumem. Vztah přesnosti (fidelity) pro  $J_K$  větší než 0.5 je dán:

$$F_k(x, y) = P\{J_k(x, y) + \xi(x, y) > 0.5\} = \phi\left(\frac{0.5 - J_k(x, y)}{\sigma}\right) \quad (5)$$

Pro  $J_K$  menší než 0.5 je vztah:

$$\begin{aligned} F_k(x, y) &= P\{J_k(x, y) + \xi(x, y) < 0.5\} \\ &= 1 - \phi\left(\frac{0.5 - J_k(x, y)}{\sigma}\right) = \phi\left(\frac{J_k(x, y) - 0.5}{\sigma}\right) \end{aligned} \quad (6)$$

$\xi$  je bílý šum s variancí  $\sigma^2$  a  $\phi$  představuje normální rozdělení pravděpodobnosti. Věrohodnost bodu si proložíme ještě přes všechny snímky použité pro rekonstrukci:

$$F(x, y) = \sum_{k=1}^N w_k F_k(x, y) = \sum_{k=1}^N w_k \phi\left(\left|\frac{0.5 - J_k(x, y)}{\sigma}\right|\right) \quad (7)$$

V tomto vztahu reprezentuje  $w_k = 2^{-k}$  důležitost pruhu, ve kterém bod leží a  $N$  představuje počet vzorů promítnutých na objekt. Na obrázku [2.4] vidíme náš zkoumaný objekt a v červených kruzích jsou ukázány právě problémy s intenzitou, která se na sádrové bystě odráží a narušuje postup. Můžeme zkoumat blízké okolí bodu a podle zjištěných intenzit přiřadit hodnotu bodu. Když se podíváme na snímek vidíme, že některé plochy v černých pruzích jsou dosti obsáhlé a mohou způsobovat chyby v algoritmu. Pokud jsou známy informace o bodech na rozhraní pruhů a víme apriorně, že první horní pruh je vždy tmavý a dále se pruhy střídají, stačí se podívat na pixely v ose kolmé na pruhy a vypočítat v kolikátém pruhu leží.



*Fig. 2.4: Skenovaný objekt (sádrová bysta J. A. Komenského)*

Samotná rekonstrukce pak vychází ze znalosti parametrů kamerového, projekčního a světového systému souřadnic. Tyto parametry zjistíme z kalibrace zařízení různými metodami. Hlavní myšlenka je transformace souřadnicových systémů. Musí platit vztahy pro souřadnicové systémy:

$$x_w = T^{-1}(x_c, x_p) \quad (8)$$

Pro kameru  $(x_c, C_c)$  a projektor  $(x_p, C_p)$  pak vztahy separátně vypadají takto:

$$x_c = C_c x_w \quad (9)$$

$$x_p = C_p x_w \quad (10)$$

Matice  $C_c$  a  $C_p$  lze rozepsat obecně jako vztah rotace, translace a proporcionální složka.

$$C_C = \alpha \begin{bmatrix} f_x & \gamma & x_C^0 \\ 0 & f_y & y_C^0 \\ 0 & 0 & 1 \end{bmatrix} [R_C \quad t_C] \quad (11)$$

$$C_P = \alpha \begin{bmatrix} f_P & 0 & x_P^0 \\ 0 & 0 & 1 \end{bmatrix} [R_P \quad t_P] \quad (12)$$

Vektory  $R_C$ ,  $t_C$  definují transformaci mezi soustavou souřadnic světa  $X_w$  a souřadnicovým systémem  $X_C$ . Stejně tak u projektoru představují vektory rotace a translace transformaci souřadnic. Parametry  $f_x$ ,  $f_y$  a  $f_P$  zjistíme kalibrací systému. Člen  $\gamma$  je faktor šikmosti os. Všechny vektory jsou však dány v homogenních souřadnicích a je možné, že neexistuje řešení pro vztahy (9) a (10) současně. Změníme tedy zápis souřadnic pro  $x_C$  a  $x_P$  na  $x_C = [w_C x_C, w_C y_C, w_C]^T$  a  $x_P = [w_P x_P, w_P]^T$ . Rozvedeme tedy zápisy (9) a (10):

$$w_C x_C = c_1 x_W$$

$$w_C y_C = c_2 x_W \quad (13)$$

$$w_C = c_3 x_W$$

a pro projektor:

$$w_P x_P = p_1 x_W$$

$$w_P = p_2 x_W \quad (14)$$

Parametry  $c_K$  představují k-tý řádek matice  $C_C$  a  $p_N$  zase představují n-tý řádek matice  $C_P$ . V obou soustavách vidíme vztah pro  $w_C$  a  $w_P$  v posledních rovnicích. Dosadíme tyto vztahy do rovnic:

$$x_C c_3 x_W = c_1 x_W$$

$$y_C c_3 x_W = c_2 x_W \quad (15)$$

$$x_P p_2 x_W = p_1 x_W$$



Z těchto vztahů v bodě (15) lze vyjádřit  $x_W$  do maticového tvaru  $Qx_W = 0$ , kde

$$Q = \begin{bmatrix} x_C c_3 - c_1 \\ y_C c_3 - c_2 \\ x_P p_2 - p_1 \end{bmatrix} \quad (16)$$

Matice  $Q$  může být dále rozdělena na dvě složky o dimenzích  $3 \times 3$  a  $3 \times 1$ :  $Q = [R, s]$ . Substituujeme  $x_W$  tak, že rozšíříme vektor o prvek  $\alpha$ . Tato úprava nijak nemění význam vektoru, když rozšíříme všechny prvky vektoru o tento nový člen. Změnili jsme tedy dimenzi vektoru přidáním jedničky. Tato změna má pouze význam pro úpravu soustavy:

$$[R, s] \begin{bmatrix} \alpha x_W \\ \alpha y_W \\ \alpha z_W \\ \alpha \end{bmatrix} = R \begin{bmatrix} \alpha x_W \\ \alpha y_W \\ \alpha z_W \end{bmatrix} + \alpha s = 0 \quad (17)$$

Parametr  $\alpha$  je možné vytknout a podělit. Bod objektu v nehomogenních světových souřadnicích  $x_W = [x_W, y_W, z_W]^T$  je pak řešením lineárního systému

$$Rx_W = -s. \quad (18)$$

Zpětná projekce je pak vyjádřena inverzí matice  $R$ :

$$x_W = -R^{-1}s \quad (19)$$

Vztah vypadá, že jsme odstranili parametry  $x_C$ ,  $y_C$  a  $x_P$ , ale musíme si připomenout, že  $R$  a  $s$  jsou funkce využívající tyto parametry. Bez těchto parametrů by nebylo možné rekonstruovat scénu. Matice inverzní k matici  $R$  existuje vždy. Jedinou výjimkou je případ, když jsou kamera a projektor vůči sobě rovnoběžné. Na obrázku [2.3] je vidět postup a vztah mezi jednotlivými kroky.

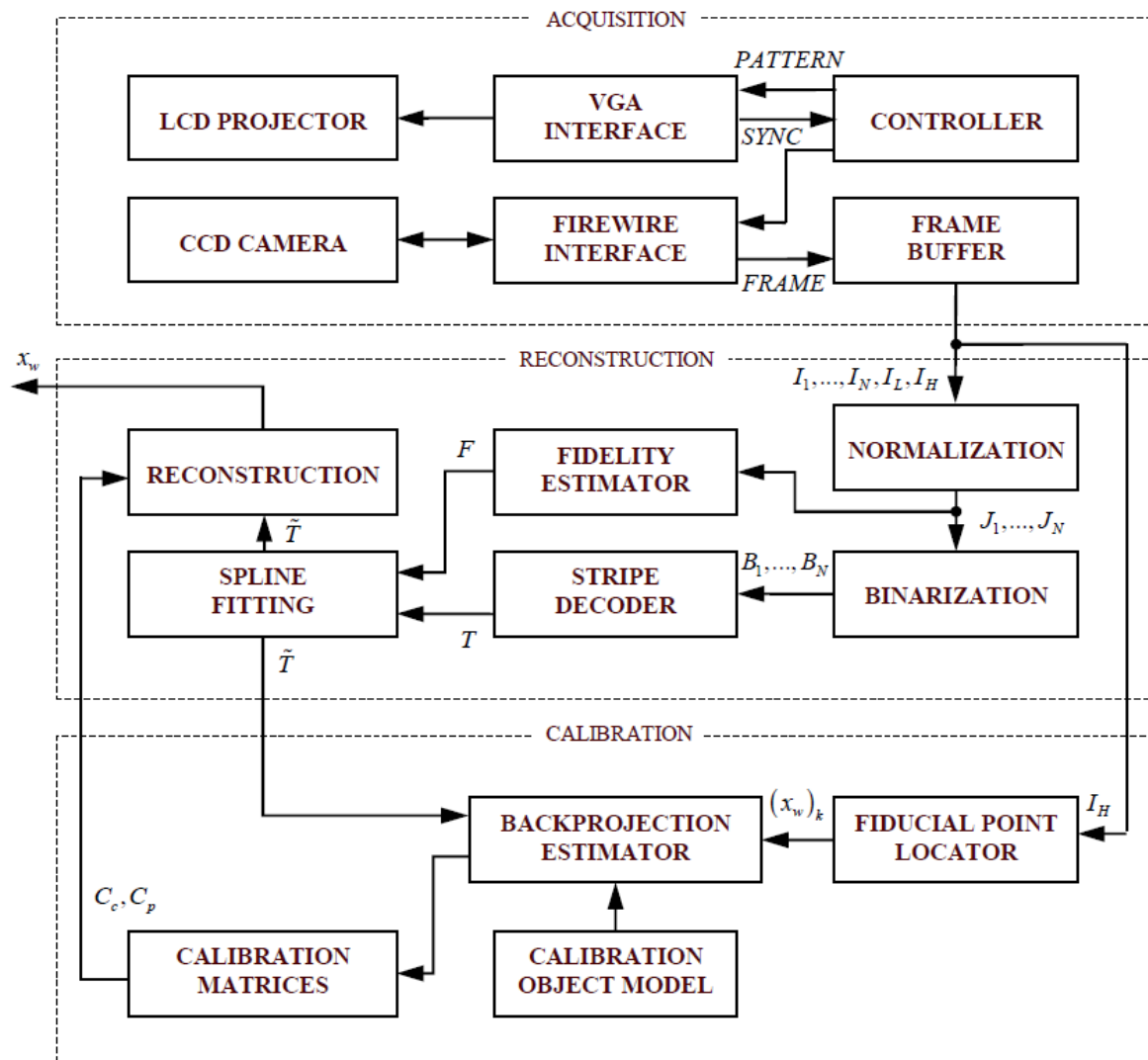


Fig. 2.5: Diagram postupu rekonstrukce v práci ([4], s. 11)

Bloky *LCD projector* a *CCD Camera* reprezentují projektor a kameru v systému. Projektor dává informaci zpět do systému pomocí snímků z kamery, tudíž ve schématu jde do tohoto bloku pouze vstup, ale nemá žádný výstup reprezentovaný. Důležitým prvkem je zde *controller*. Tento blok je v reálu reprezentován uživatelem nebo algoritmem, který postupně zadává informaci o vzoru do projektoru a určuje časový okamžik záběru snímků. *Camera* tedy jako vstup má informaci, kdy snímat objekt a výstupem bloku je snímek, který pak kamera posílá

zpátky do bloku vyrovnávací paměti (*buffer*), ze které si pak snímky bereme pro kalibraci a rekonstrukci. Tyto bloky doposud popsané lze substituovat jako subsystém, který slouží jako generátor vstupů. Ve schématu je tento generátor nazván *acquisition* (získávání). Druhý subsystém je rekonstrukce. Vstupem jsou všechny snímky s promítnutým vzorem a snímek při plném osvětlení a neosvětlený. Snímky  $I_1, I_2, \dots, I_N, I_L$  a  $I_H$  se nejprve v prvním bloku normalizují vztahem popsaným ve vztahu (3) a získáme normalizované snímky  $J_1, J_2, \dots, J_N$ . Každý bod v tomto snímku má tedy přiřazenou hodnotu v rozmezí  $\langle 0,1 \rangle$ .  $J_1, J_2, \dots, J_N$  pak slouží jako vstup pro blok, který provádí binarizaci a přiřazuje bodu přesnou hodnotu 1 nebo 0. Tyto nové snímky označíme jako  $B_1, B_2, \dots, B_N$ .  $J_1, J_2, \dots, J_N$  jsou vstupem i pro blok, který vypočítává pravděpodobnost správného přiřazení hodnoty k bodu. Tento blok je popsán ve vztazích (5),(6) a (7).  $B_1, B_2, \dots, B_N$  nyní poslouží k určení identifikace pruhů (*stripe ID*) v bloku *stripe decoder*. Stripe id  $T$  nám udává souřadnici v souřadnicové soustavě projektoru. Vlivem chyb ve snímcích však může nastat že pruhy  $T$  nejsou hladké. Blok *spline fitting* tedy pruhy zahladí a vytvoří upravený pruhy  $\tilde{T}$ . Tyto upravené pruhy pak používáme pro samotnou rekonstrukci v bloku *reconstruction* a dává nám výstup celého systému, kterým je bod  $x_W$  v prostorových souřadnicích. Tato metoda ve schématu však používá ke kalibraci metodu zpětné projekce. Subsystém kalibrace zde v této metodě využívá snímků objektu, který má exaktně dané parametry a je na něm vyznačen soubor bodů, které také musí na objektu být umístěny s velkou přesností. Tyto body nazýváme výchozími body (*fiducial points*). Na *figuře 2.4* je znázorněn použitý kalibrační objekt s vyznačenými výchozími body. Vstupy do kalibrace jsou tvořeny snímkem osvětleného kalibračního objektu a a vyhlazených pruhů  $\tilde{T}$ . Snímek s objektem nejprve projde algoritmem nalézajícím výchozí body (*fiducial point locator*) a pak společně s vyhlazenými pruhy  $\tilde{T}$  a matematickým

modelem kalibračního objektu se provede zpětná projekce v bloku *backprojection*. Po těchto krocích je možné určit parametry matic  $C_c$  a  $C_p$ , které následně aplikujeme v rekonstrukci. Tato metoda nám tedy praví, že není potřeba provádět kalibraci systému separátně od rekonstrukce, ale je potřeba pro první spuštění skenovat kalibrační těleso, jehož matematický model známe, nebo je možné jej zjistit a zároveň, aby těleso bylo možné jej vyrobit. Čtyřstěn použitý zde je dobrým příkladem.

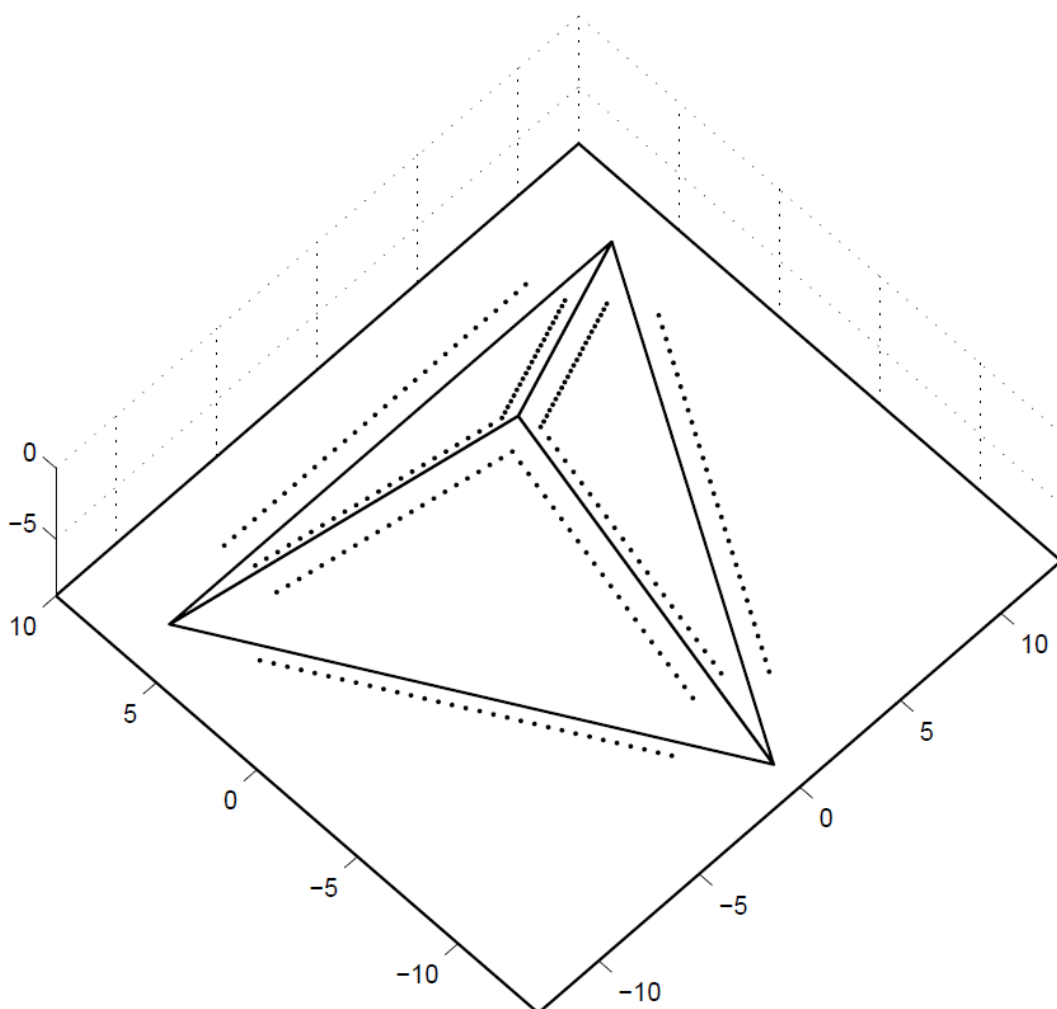


Fig. 2.6: Kalibrační objekt použitý v práci [5]

### 2.2.3 Realizace

Náš systém jsme sestavili pomocí zrcadlového fotoaparátu značky Olympus, projektoru značky Canon a notebooku. Projektor byl připojen na notebook pomocí VGA kabelu. Projektor sloužil jako druhý monitor a na notebooku jsme pak v režimu *Celá obrazovka* v prohlížeči obrázků postupně otevírali jeden snímek po druhém a mezitím jsme ručně pomocí fotoaparátu snímali scénu. V průběhu jsme mohli kontrolovat snímky na displeji kamery. Celý soubor dat jsme pak z kamery přenesli do notebooku ke zpracování přes SD kartu. Další postup už se prováděl pouze na počítači v prostředí MatLab.



*Fig. 2.7: Naše sestava v reálném světě*



*Fig. 2.8: Pohled na soustavu jak promítá vzor na objekt*

V MatLabu jsme pak využívali pro kalibraci toolbox od Jean-Yves Bouget. Toolbox má několik podmínek pro používání. První podmínkou je kontrola, zda je toolbox podporovaný používanou verzí MatLabu. Dále je pak nutné mít správně pojmenovaná data a složku s nimi mít přidanou do vstupů MatLabu. Po snímání je tedy důležité zkontrolování snímků. Fotoaparát podával kvalitní snímky, které neukazovali známky kruhovitěho zakřivení, takže jsme se zaměřovali pouze na ostrost snímků a pokud aplikovaný filtr v kameře nezpůsoboval přílišný odraz. Po profiltrování snímků je pak doporučeno je vhodně pojmenovat. Toolbox název snímků dělí na prefix a sufix. Pro kalibraci kamery je třeba vhodné použít prefix *cam*. Prefix od sufixu nemusí být od sufixu oddělen dalším znakem. Jména snímků mohou tedy vypadat takto:

cam001, cam002,...cam999 nebo cam1,cam2,...cam999

Poté co jsou všechny snímky správně pojmenovány spustíme toolbox. Nejprve ho najdeme mezi ostatními toolboxy ve složce MatLabu a kliknutím na jeho složku dáme *Add to path*. Nyní je možné použít příkaz *calib*. Spustí se uživatelské rozhraní (GUI), které nás nejprve vyzve ke zvolení metody ukládání snímků do paměti. Tato volba nijak neovlivňuje výsledky kalibrace, ale u některých počítačů je vhodné zvolit efektivnější metodu, kdy je v každém okamžiku kalibrace v paměti uložen jen jeden snímek. Dalším krokem je tedy zadání snímků, které použijeme pro kalibraci.

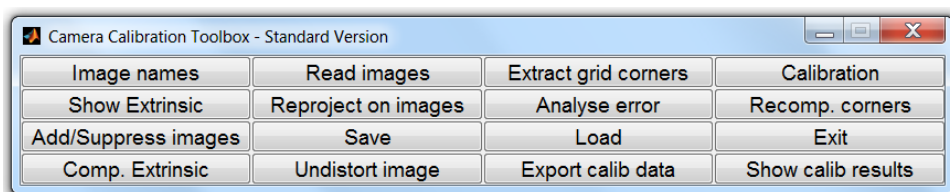
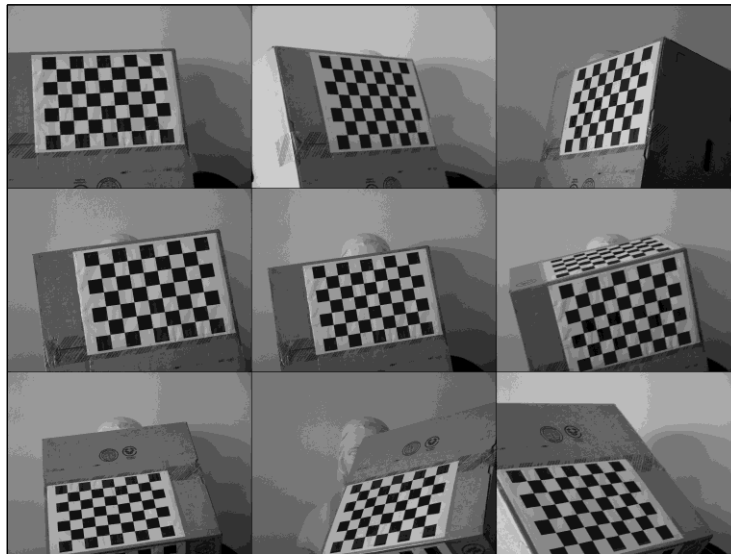


Fig. 2.9: Uživatelské rozhraní pro kalibraci kamery

Nyní musí být v MatLabu zvolena složka obsahující kalibrační snímky. Pokud je tato složka vybrána pak kliknutím na tlačítko *Image names* (Názvy snímků) nám do *Command window* (příkazové okno) v MatLabu vypíše seznam všechny snímky všech formátů. Zeptá se nás, abychom zadali prefix snímků pro kalibraci. Napíšeme příslušnou předponu a stisknutím *Enter* volbu potvrdíme. V dalším kroku se nás zeptá na formát snímků. Každý formát (jpg, bmp,...) má příslušnou zkratku. Seznam zkratk pro formáty se vypíše před zadáním. Po potvrzení formátu jsou snímky načteny a připraveny.

Calibration images



*Fig. 2.10: Načtené snímky*

Snímky kalibrace musí obsahovat šachovnici. My jsme připevnili vytisknutou šachovnici na kartonovou krabici a postavili ji v libovolném pohledu před kameru. Snímky by neměli být stejné. Je potřeba šachovnici snímat pod různými úhly. Abychom mohli provést kalibraci je potřeba znát velikost strany jednoho čtverečku v milimetrech. Tlačítkem *Extract grid corners* (Vytáhni rohy mřížky) se postupně v MatLabu otvírají jednotlivé snímky. Ručně pak klikáme na body jako ve figurě 2.11.



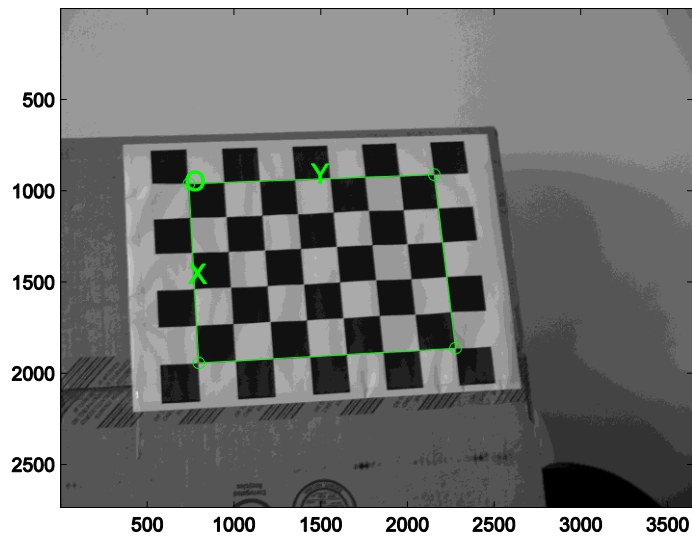


Fig. 2.11: Ukázka zadávání rohů

Toto provedeme pro všechny snímky a nakonec vidíme jak přesně jsme postupovali na finálním snímku, kde vidíme jak se podařilo nalézt všechny body mezi každou čtveřicí čtverců. Nyní je možné kliknout na tlačítko *Calibration*. V příkazovém okně se nám pak vypíše výsledky kalibrace. Do složky se snímky se pak vytvoří soubor *calib\_data.mat* obsahující hodnoty kalibrace a je možné je později načíst. Podobným způsobem postupujeme při kalibraci projektoru. Zadáním příkazu *cam\_proj\_gui* se otevře nové GUI (Fig. 2.12).

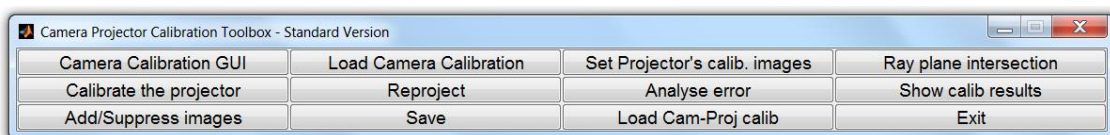
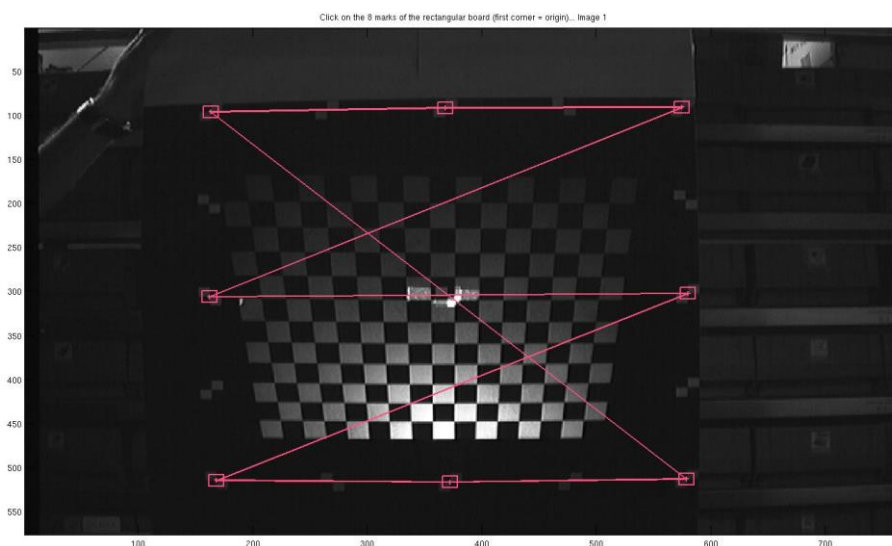


Fig. 2.12: GUI pro kalibraci projektoru

Pokud bychom neměli kalibrovanou kameru bylo by možné kliknutím na tlačítko *Camera Calibration GUI* přepnout do předchozího postupu. Hodnoty kalibrace máme již ve zmíněném *calib\_data.mat* souboru. Snímky pro kalibraci projektoru máme uloženy ve stejné složce a není

nutné přepínat mezi složkami a kopírováním onoho *mat* souboru. Stačí tedy načíst hodnoty kalibrace kamery tlačítkem *Load Camera Calibration*. Postupujeme stejně jako v případě kamery. Jediným rozdílem je, že budeme kromě šachovnice zkoumat i plochu na kterou je šachovnice promítána. Vyznačíme si tedy na této ploše jasně viditelné značky. Těchto značek bude osm. Čtyři z osmi budou v rozích vymezeného čtverce a zbylé čtyři budou v polovině mezi každými dvěma body.



*Fig. 2.13: Ukázka promítací plochy a vyznačených bodů*

Další postup je totožný s kalibrací kamery použitím tohoto toolboxu a výsledky kalibrace se uloží do souboru *calib\_results.mat*. Pro další práci s rekonstrukcí jsme si připravili algoritmus v MatLabu přiložený v příloze, který nám generuje podle zadaného rozlišení binární vzory. Vygenerované vzory jsou ve formátu *bmp* (bit mapový formát). Důvodem stvoření tohoto generátoru bylo dosažení potřebného počtu vzorů při daném problému a aby při ztrátě bylo možné snadno vygenerovat nové přesné vzory.

### 3 Kalibrace

Kalibrace je prvním a nejdůležitějším krokem celého procesu rekonstruování objektů v počítačovém vidění. Právě v tomto kroku se určují parametry a vlastnosti mezi kamerami a projektory. V druhé kapitole jsme již ukázali, že pro výpočet prostorových souřadnic jsou důležité určité vlastnosti zařízení. Parametry  $f_x$ ,  $f_y$  a  $f_p$  v maticích  $C_C$  a  $C_P$  jsou ony vlastnosti zařízení. Pro kameru jsou  $f_x$  a  $f_y$  ohniskové vzdálenosti pro obě dimenze 2D prostoru. U projektoru je  $f_p$  také ohnisková vzdálenost, ale na rozdíl od kamer má projektor pouze jeden tento parametr. Kalibračních metod rozdělujeme fotogrammetrické (photogrammetric calibration) a na samo–kalibrační (self-calibration). První ze zmíněných metod provádí kalibraci sledováním kalibračního objektu, jehož vlastnosti v prostoru jsou s velkou přesností známy. Druhý přístup nevyžaduje kalibrační objekt, ale stačí s kamerou pohybovat a snímat statickou scénu. Za předpokladu, že vnitřní parametry kamery jsou neměnné, je pak z pouhých tří snímků obdržet vnitřní a vnější parametry. Je však nutné, aby byly k nalezení korespondenční body mezi jednotlivými snímky. Vnitřními parametry jsou ohniskové vzdálenosti zařízení a vnějšími parametry nazýváme vektory popisující rotaci a translaci.

#### 3.1 Kalibrace kamery

U kamer je třeba řešit před kalibrací i linearizaci obrazu pro případ, že čočka kamery je nedokonalá a nastává radiální zkreslení nebo také jev rybiho oka. Snímek je zakřiven do kruhů, jako když se člověk podívá kukátkem ve dveřích. Takovéto snímky je třeba linearizovat před samotnou kalibrací. Budeme uvažovat o snímku, který by nebyl zakřivený a body by měli souřadnice  $[x, y]$  a pak vezmeme snímek zakřivený ve kterém

máme body  $[\check{x}, \check{y}]$ . Můžeme pak zavést vztahy mezi body v zakřiveném a nezakřiveném snímku:

$$\begin{aligned}\check{x} &= x + x[k_1(x^2 + y^2) + k_2(x^2 + y^2)^2] \\ \check{y} &= y + y[k_1(x^2 + y^2) + k_2(x^2 + y^2)^2]\end{aligned}\quad (20)$$

Parametry  $k_1$  a  $k_2$  jsou koeficienty radiálního zakřivení.



*Fig. 3.1: Ukázka zakřivení snímku – problém rybího oka*

Bod v ploše je definován dvěma souřadnicemi a je zapisován  $b = [x, y]^T$ . V prostoru má bod souřadnice tři a je zapisován ve tvaru  $B = [X, Y, Z]^T$ . Pro vztahy mezi maticemi je však nutné zavést do zápisů rozšíření, které nemá žádný fyzikální význam pro souřadnice, ale umožňuje výpočty. V 2D prostoru bude nový zápis  $\tilde{b} = [x, y, 1]^T$  a pro bod v prostoru bude  $\tilde{B} = [X, Y, Z, 1]^T$ . Pro tyto rozšířené zápisy pak platí vztah mezi plochou a prostorem:

$$s\tilde{b} = C_c [R \ t] \tilde{B} \quad (21)$$

Matice  $C_C$  jsou vnitřní vlastnosti kamery a matice  $[R \ t]$  jsou vnější parametry. Prvek  $s$  je skalár.

### 3.1.1 Zhangova metoda

Metoda kalibrace zde vychází z analyzování vzoru na ploše, která je snímána z různých pohledů. Postupujeme tak, že vytiskneme vzor šachovnice a tu připevníme na rovný povrch. Dále pak stačí nasnímat tento vzor v několika úhlech a vzdálenostech. Snímky můžeme pořizovat pohybem kamery, ale pokud chceme zachovat vztah mezi kamerou a dalším zařízením v prostoru je lepší pohybovat vzorem v prostoru před kamerou na rovné pevné ploše. Tato metoda má jen málo pravidel na sledování a při své jednoduchosti je přesná. Kalibrace touto metodou je možná již dvou snímci, ale je doporučeno mít více snímků pro lepší výsledky. Pro  $n \geq 3$ , kde  $n$  je počet snímků, dostáváme již velice přesné výsledky. Vyjdeme ze vztahu (21) a bod  $\tilde{B}$  leží v rovině  $Z$ . Upravíme tedy vztah:

$$s \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = C_C [r_1 \ r_2 \ r_3 \ t] \begin{bmatrix} X \\ Y \\ 0 \\ 1 \end{bmatrix} = C_C [r_1 \ r_2 \ t] \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} \quad (22)$$

Prvky  $r_i$  jsou jednotlivé sloupce rotační matice  $R$ . Prvek  $r_3$  odstraněním z-souřadnice ztratil význam a mohli jsme jej opomenout. Nyní si zavedeme substituci homografie v našem upraveném vztahu  $H = C_C [r_1 \ r_2 \ t]$  a dostaneme vztah mezi body  $\tilde{m}$  a  $\tilde{M}$  pomocí homografie.

$$s\tilde{m} = H\tilde{M} \quad (23)$$

Matice homografie bude mít dimenze 3x3, aby vyhovovala všem vztahům popsaným výše. Řešením homografie pak získáme parametry soustavy.

### 3.2 Kalibrace projektoru

Při kalibrování kamery jsme vycházeli z apriorní znalosti parametrů kalibračního tělesa nebo vzoru. U projektoru musíme promítat nějaký vzor použitelný pro kalibraci a snažit se ho transformovat do souřadnicového systému kamery. Postup pro kalibraci projektoru je podobný jako u Zhangovi metody. Projektor může být považován za inverzní kameru, jelikož obraz promítá a nezachycuje. S touto myšlenkou lze kalibrovat projektor kamerovou kalibrační metodou, pokud je možné nalézt hraniční body vzoru na snímcích. Kalibraci celého systému jsme prováděli v prostředí MatLab pomocí toolboxu od Jean-Yves Bouget.

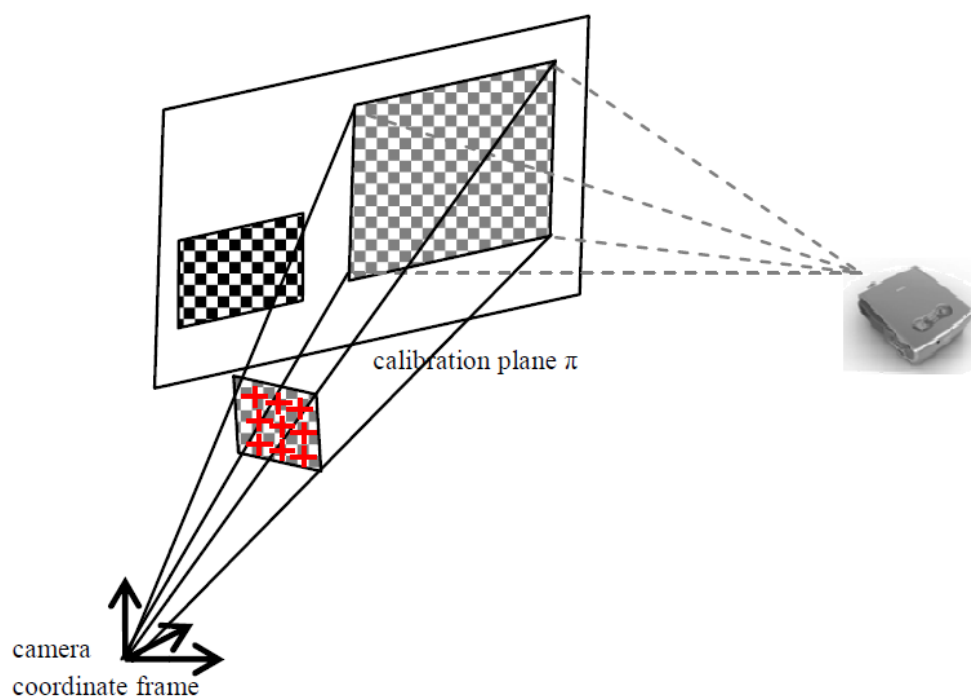


Fig. 3.2: Pohled na systém kamera-projektor

Celý postup kalibrace projektoru závisí tedy na nalezení průsečíků paprsků procházející mezními body promítaného vzoru a roviny na kterou je vzor promítán.

### 3.3 Výsledky kalibrace

Pro kalibraci kamery jsme použili sadu devíti snímků.

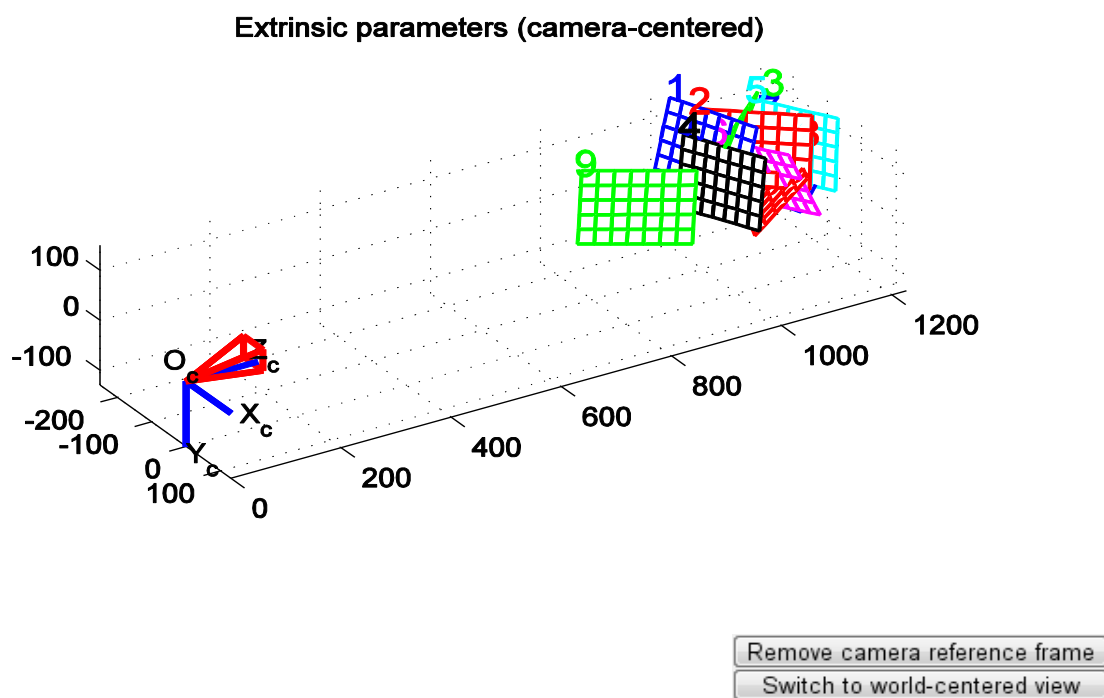


Fig. 3.3: Zrekonstruování polohy vzoru při kalibraci

```

Initialization of the intrinsic parameters - Number of images: 9

Calibration parameters after initialization:

Focal Length:      fc = [ 8552.66025  8552.66025 ]
Principal point:   cc = [ 1823.50000  1367.50000 ]
Skew:              alpha_c = [ 0.00000 ] => angle of pixel = 90.00000 degrees
Distortion:        kc = [ 0.00000  0.00000  0.00000  0.00000  0.00000 ]

Main calibration optimization procedure - Number of images: 9
Gradient descent iterations: 1...2...3...4...5...6...7...8...9...10...11...12...13...14...15...16...17...18...19...20...21...22...23...2
Estimation of uncertainties...done

Calibration results after optimization (with uncertainties):

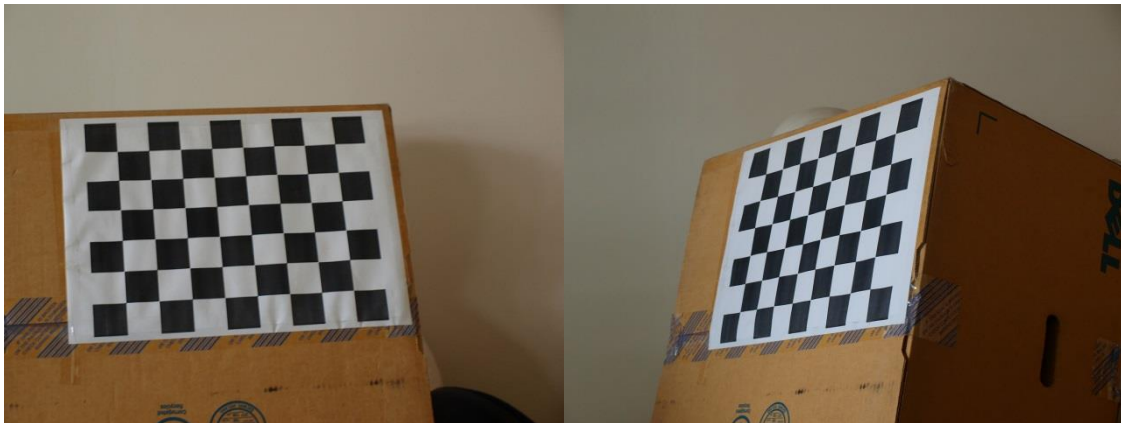
Focal Length:      fc = [ 8708.54939  8812.21324 ] d'ž [ 192.97527  220.62225 ]
Principal point:   cc = [ 1902.53406  671.92087 ] d'ž [ 199.57302  288.91413 ]
Skew:              alpha_c = [ 0.00000 ] d'ž [ 0.00000 ] => angle of pixel axes = 90.00000 d'ž 0.00000 degrees
Distortion:        kc = [ -0.24546  2.92805  -0.01767  0.00285  0.00000 ] d'ž [ 0.17025  2.36539  0.01586  0.00593  0.00000 ]
Pixel error:       err = [ 1.56108  1.90553 ]

Note: The numerical errors are approximately three times the standard deviations (for reference).

```

*Fig. 3.4: Výsledky kalibrace kamery v matlabu*

Všechny snímky pro kalibraci jsou přiloženy na konci této práce. Pro ukázkou dva snímky vypadaly takto:



*Fig. 3.5: Kalibrační snímky (Kalibrační vzor připevněný na pevnou pohyblivou podložku)*

Samotná kalibrace projektoru se nám bohužel nepodařila taková, aby dávala uspokojivé výsledky v ideálním případě by pak rekonstrukce promítaného vzoru byl podobný jako rekonstrukce šachovnice v obrázku [3.3] a vypadal by například takto:



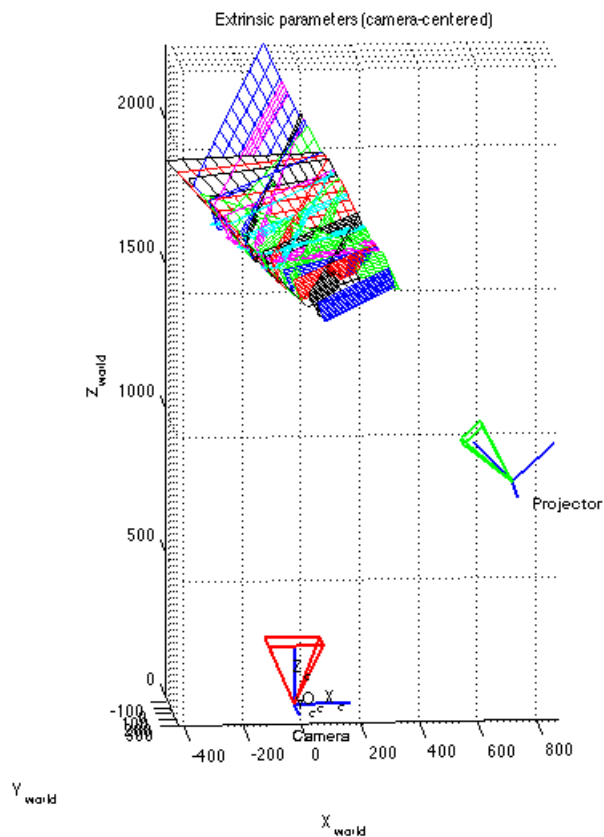


Fig. 3.6: Ukázka kalibrace projektoru z návodu pro *Cam\_Proj\_Calib* toolbox

Postupovali jsme sice podle instrukcí, ale reálně výsledný výpočet mezi kamerou a projektorem dával nesmyslné relace. V druhé kapitole vidíme jak byli zařízení vůči sobě umístěné, ale výsledek udával, že by projektor měl být až dva a půl metru za kamerou a promítané vzory by se museli promítat do další místnosti za stěnou, jelikož ty vycházeli dál než by měli. Bez správné kalibrace však jsme nemohli přistoupit k dalšímu kroku.

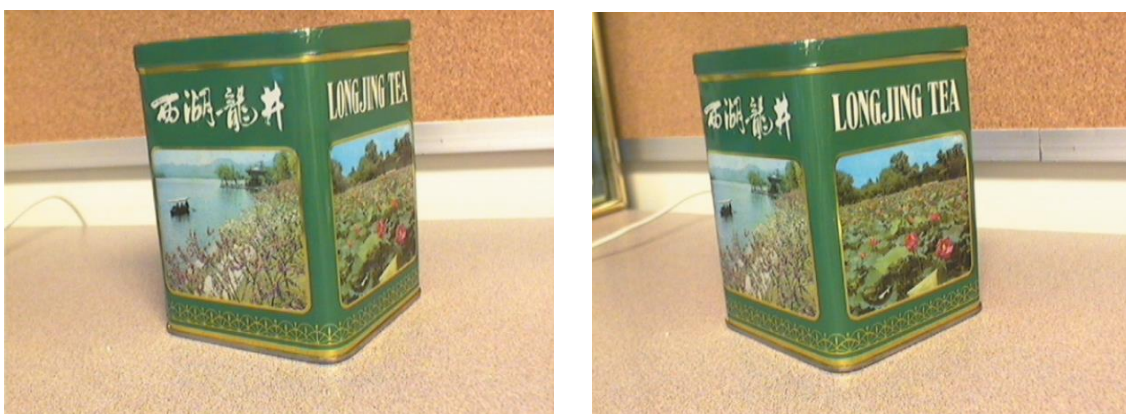
## 4 Shrnutí

Definovali jsme si problematiku počítačového vidění a různá specifika této oblasti. Ukázali jsme vlastnosti a postup rekonstrukce za použití metody strukturovaného světla Binárními vzory vygenerované v prostředí MatLab. Toolbox Jean-Yves Bougeta je kvalitním nástrojem a je dostupný zadarmo na internetu. Problémem tohoto toolboxu je však to, že jeho snadnou dostupností byl již několikrát modifikován různými lidmi pro své specifické případy a toolbox potřebuje hlubší znalost jeho struktury pro snazší zpracování analýzy. Základní toolbox však pro kalibraci kamery je nezměněn a Zhangova metoda splnila naše očekávání na kalibraci kamery a poskytla žádané výsledky. Pro projektor je třeba nalézt správný kalibrační postup. V dalším výzkumu se zaměříme na kalibraci projektoru hlouběji a sestavíme celý aparát 3D skeneru.

## 5 Ukázky a přílohy

Snímky pořízené pro kalibraci projektoru a pro práci rekonstrukce jsou přiloženy na CD, které je součástí této práce. Podíváme se rekonstrukce objektů různými metodami:

Stereovize (dvě kamery):



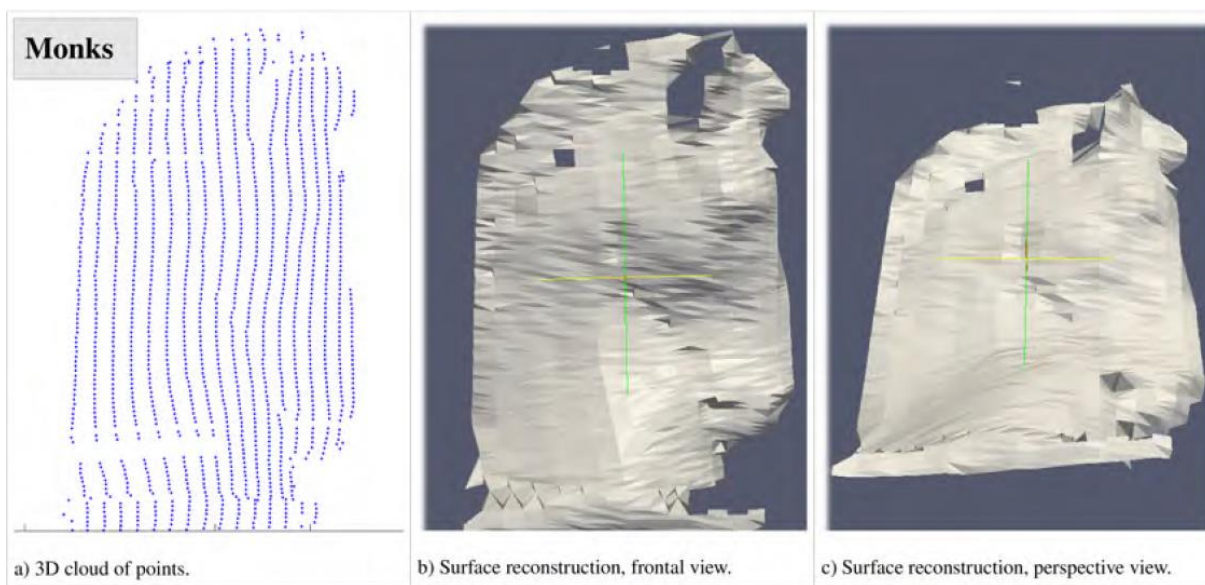
*Fig. 5.1: Dva pohledy na jeden objekt pro stereovizi*



*Fig. 5.2: Rekonstrukce objektu ze snímků v 5.1 s texturou*

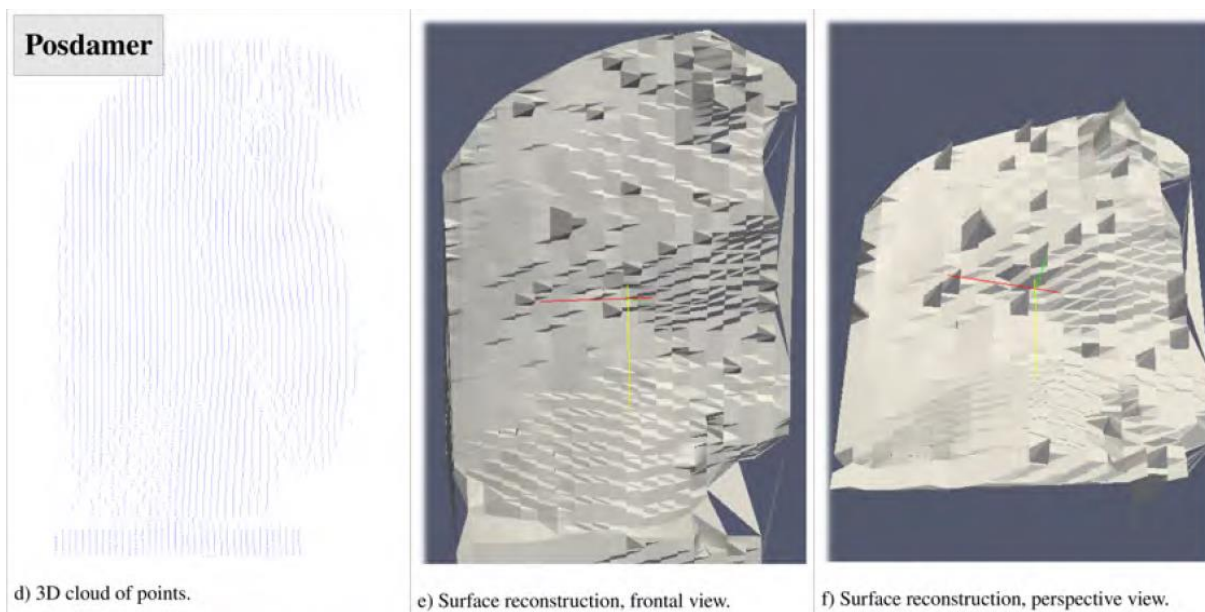
Porovnání metod strukturovaného světla:

Barevný vzor (De Bruijn viz Fig. 2.1) – rekonstrukce Monkovo metodou:



*Fig. 5.3: Výsledky Monkovo metody*

Binární vzor pomocí Posdamerovo metody:



*Fig. 5.4: Výsledky Posdamerovo metody*

Vzor s fázovým posunem pomocí Guhringovi metody

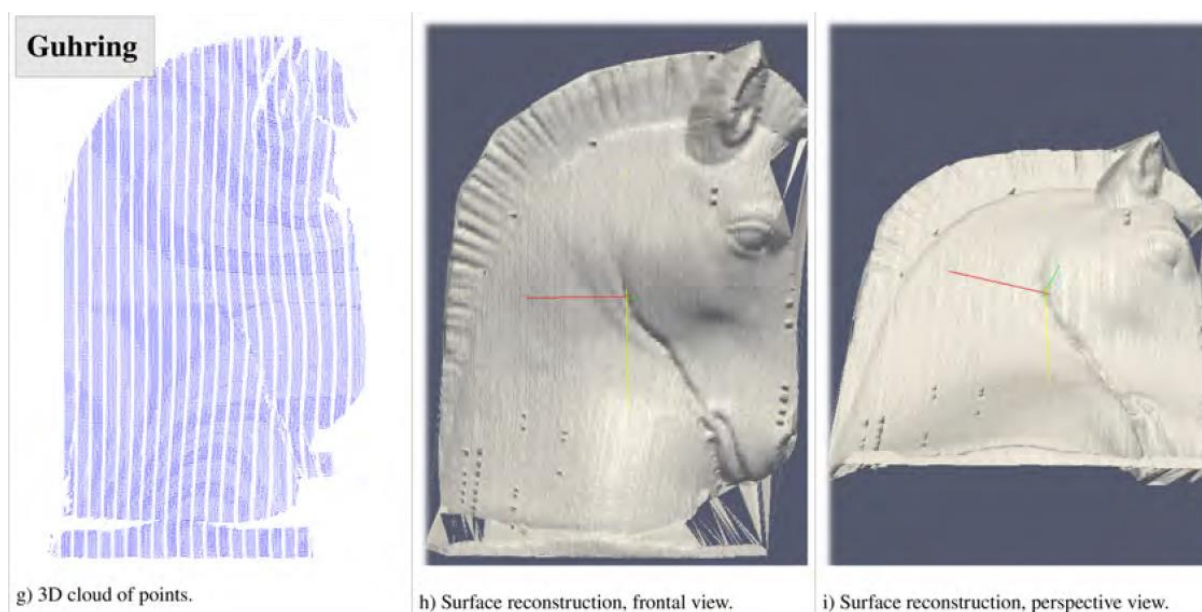


Fig. 5.5: Výsledky Guhringovi metody

Vidíme, že každý přístup má různé hustoty bodů v mapě a že Guhringova metoda pro rekonstrukci objektu má velice dobré výsledky v porovnání s Posdamerem a Monksem.

V metodě popsanou v [5] prováděli rekonstrukci lidské tváře a dosažené výsledky jsou velice přesné:

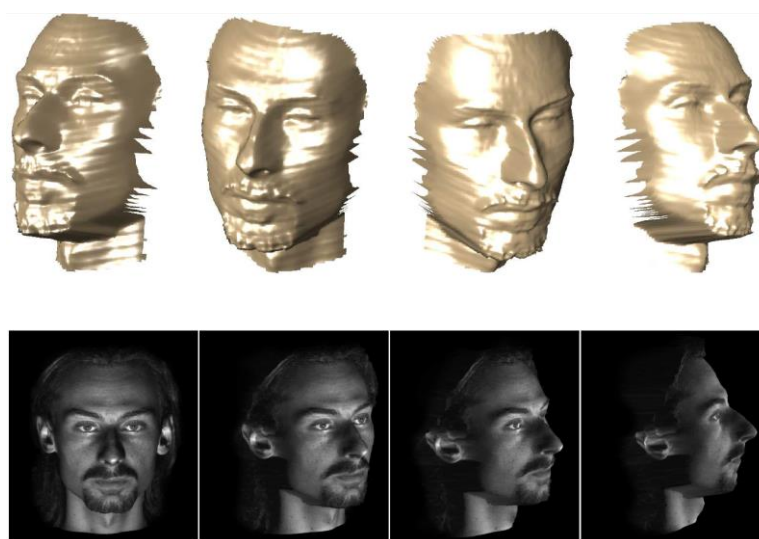


Fig. 5.6: Výsledky rekonstrukce binárním vzorem ([5], s. 19)

## Literatura

- [1] Navarro, S.,F.; 3D reconstruction of object shae using structured light. *Universitat de Girona*, 2009.
- [2] Bouget, J.; Camera calibration toolbox for matlab ([http://www.vision.caltech.edu/bougetj/calib\\_doc/](http://www.vision.caltech.edu/bougetj/calib_doc/)).
- [3] Navarro, S.,F.; One-shot pattern recognition for dense and accurate 3D reconstruction in structured light. *Universitat de Girona*, 2012.
- [4] Zhang, Z.; A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(11): 1330-1334, 2000, 1998.
- [5] Bronstein, A.,M.; Bronstein, M., M., Gordon, E., Kimmel, R.; High-resolution structured light range scanner with automatic calibration, *Israel Institute of Technology*, 2003.
- [6] Falcao, G., Hurtos, N., Massich, J.; Plane-based calibration of a projector-camera system. *VIBOT Master*, 2008.
- [7] Hartley, H., Zisserman, A.; Multiple view geometry in computer vision – second edition. *vydáno na univerzitě v Cambridge, Anglie*, 2003.
- [8] Geng, J.; Structured-light 3D surface imaging: a tutorial. *IEEE Intelligent Transportation System Society, 11001 Sugarbush Terrace, Rockville, Maryland 20852, USA*, 2010.
- [9] Forster, F.; A high-resolution and high accuracy real-time 3D sensor based on structured light. *Proceedings of the Third International Symposium on 3D Data Processing, Visualisation, and Transmission (3DPVT06), IEEE Computer Society Washington, DC, USA*, 2006.

- [10] Posdamer, J., Altschuler, M.; Surface measurement by space-encoded projected beam systems, *Proceedings of the Third International Symposium on 3D Data processing, Visualizaion, and Transmission*, Computer Society Washington, 1982.
- [11] Pages, J., Salvi, J., Forest, J.; A new optimised De Bruijn coding strategy for structured light patterns. *Proceedings of the 17th international Conference, Pattern Recognition*, 2004.