

Oponentní posudek na doktorskou disertační práci.

Autor disertační práce:	Ing. Milan Košťák
Název disertační práce:	Využití umělé inteligence v oblasti rozšířené reality.
Školitel:	prof. RNDr., PhDr., Antonín Slabý, CSc.
Studijní program:	Aplikovaná informatika
Studijní obor:	Aplikovaná informatika
Školící pracoviště:	Katedra informatiky a kvantitativních metod, FIM UHK

Předložená disertační práce (DP) má 97 stran textu rozděleného do deseti kapitol (97 stran), 5 stran příloh, seznam publikačních aktivit autora (1strana) a seznam 114 použitých literárních zdrojů. *Úvodní kapitola* obsahuje vhléd do problematiky rozšířené reality. Zavádí pojem rozšířené reality a obrazové značky tzv “marker”, kdy Marker je předem definovaným tvarem (speciální vizuální značkou – referenčním bodem), který je ve scéně snadno detekovatelný a podle něho lze zjistit pozici, orientaci a další nezbytné informace pro zajištění přesné vzájemné interakce reálného a virtuálního prostředí. *V druhé kapitole* je představen cíl disertační práce rozdělený do dvou podcílů. **První z nich je návrh systému obrazových značek (tzv. „markerů“), které jsou díky své jednoduchosti přesně zaměřeny na využití v aplikacích rozšířené reality, to jest návrh markeru, který umožní detekci konkrétního místa ve snímaném obrazu v systému rozšířené reality. Druhým podcílem je prozkoumání možností detekce markerů navržených v rámci prvního hlavního cíle práce.** Navrhovaný marker musí splňovat několik kritérií, která jsou definována jako dílčí cíle při návrhu, který je prvním hlavním cílem práce. Kritéria jsou následující:

- schopnost lokalizace markeru (preferovaně jeho středu),
- schopnost detekce orientace jako úhlu rotace v intervalu 0° až 360°,
- a schopnost získání relativní velikosti pomocí detekce celého markeru.

Třetí kapitola se zabývá metodikou zpracování DP, evaluačními metrikami testování, měřením času zpracování algoritmů a použitím různých datových sad pro trénování neuronové sítě. Všechny datové sady byly vždy náhodně rozděleny v poměru 7:2:1 na trnovací, validační a testovací sadu. Výjimkou byla pouze finální čtvrtá sada, kde rozdělení bylo v poměru 7:1:2, aby byl zajištěn dostatečný a reprezentativní počet pro finální testování. *Čtvrtá kapitola* je věnována rešerši literatury věnující se dané problematice. Rešerši autor prováděl pro roky 1996 až 2020. Toto období si rozdělil na úseky 1996-2008, 2009-2015, 2016-2020. V dílčím závěru této kapitoly autor uvádí, že většina představených markerů je podobného typu jako dnes všudypřítomné QR kódy, jejichž cílem je uložit ve scéně co nejvíce (textových) informací, které jsou snadno strojově čitelné. *V páté kapitole*

se autor zabývá procesem návrhu marker. Kapitola je rozdělena do deseti podkapitol, která každá popisuje průběžné výsledky kontinuálního vývoje a myšlenkových pochodů, které vedly k návrhu finálního tvaru markeru, který má dobré výsledky při detekci a splňuje všechna tři definovaná kritéria uvedená v druhé kapitole. *Šestá kapitola* uvádí shrnutí výsledků návrhu markeru. Výsledkem je navržený marker „T-křížek“, který umožňuje spolehlivou detekci v rozmazaném obrazu a zároveň i možnost detekce orientace. *Sedmá kapitola* se zabývá způsoby detekce navrženého markeru. V kapitole je nejprve popsáno využití neuronových sítí pro tvorbu detektoru, dále jsou pak popsány vlastní algoritmy pro detekci markeru v obrazu. Tyto algoritmy využívají programovatelné shadery grafických karet. *Osmá kapitola* se věnuje získání informací z detekovaného markeru. Po získání souřadnic bounding boxu je nutné udělat výřez z původního obrázku a provést sérii operací, které zjistí potřebné výstupní informace. Operace lze rozdělit do dvou skupin: zpracování obrazu pro získání informací o liniích markeru a následné využití linií markeru pro výpočet informací o středu, orientaci a velikosti. *Devátá kapitola* shrnuje výsledky DP. Marker byl navrhován iterativním způsobem postupných testů, kdy bylo zkoumáno, jak se jednotlivé zvažované tvary chovají z hlediska možností detekce. V průběhu iterací byl nakonec navržen i tvar „T-křížek“, který měl jedny z nejlepších výsledků a zároveň splňuje všechna kritéria na něj kladená v rámci cílů práce:

- Schopnost lokalizace markeru – tvar má střed, který je dobře detekovatelný, jedná se o průsečík tří čar jdoucích ze tří rohů obdélníku.
- Detekce úhlu rotace v rozlišení 360° relativně k pozorovateli – tvar obsahuje tři linie, které umožňují detekci popsané rotace.
- Detekce celého markeru pro získání relativní velikosti – základem tvaru je vyplněná obdélníková oblast, která je využitelná pro detekci velikosti.

Splnění všech tří vymezených kritérií bylo nezbytné dodržet. Dále byla definována dvě doplňková kritéria, které jsou také dílčími cíli prvního hlavního cíle práce:

- jednoduchý tvar markeru včetně zajištění možnosti kreslení rukou,
- a vzor markeru bez vazby na jeho konkrétní barvu

Desátá, závěrečná kapitola zhodnocuje vytvoření markeru, který je vhodný pro aplikace využívající principů rozšířené reality.

Dále se vyjádřím k:

- Aktuálnosti zvoleného tématu disertační práce.
- Splnění sledovaného cíle.
- Zvoleným metodám zpracování disertační práce.
- Přínosům DP pro praxi a další rozvoj vědy a techniky
- K formální a jazykové úrovni disertační práce
- Úrovni publikační činnosti doktoranda a kvalitě použitých informačních zdrojů

Aktuálnost zvoleného tématu disertační práce.

Zvolené téma je aktuální, zabývá se možnostmi využití strojového učení v oblasti rozšířené reality. Soustřeďuje se zejména na aplikaci umělých konvolučních neuronových sítí (CNN), které nacházejí bohaté využití v počítačovém vidění.

Splnění sledovaného cíle.

Podle mého názoru splnění hlavního cíle disertační práce je návrh markeru, zde konkrétně „T-křížek“, který umožní detekci konkrétního místa ve snímaném obrazu v systému rozšířené reality. Druhým podcílem je prozkoumání možností detekce markerů navržených v rámci prvního hlavního cíle práce. Podrobněji viz devátá kapitola DP. Závěr: **Oba cíle byly splněny.**

Zvolené metody zpracování disertační práce.

Při zpracování DP byly použity evaluační metriky testování, měření času zpracování algoritmů. Dále byly použity různé datové sady pro trénování neuronové sítě. Všechny datové sady byly vždy náhodně rozděleny v poměru 7:2:1 na trénovací, validační a testovací sadu. Výjimkou byla pouze finální čtvrtá sada, kde rozdělení bylo v poměru 7:1:2, aby byl zajištěn dostatečný a reprezentativní počet pro finální testování.

Přínosy DP pro praxi a další rozvoj vědy a techniky.

Přínosy DP pro praxi a další rozvoj vědy a techniky můžeme spatřit právě ve splnění cílů disertační práce.

Formální a jazykové úrovni disertační práce.

K formální a jazykové úrovni disertační práce nemám zásadních připomínek. DP se nevyhnula drobným překlepům, mnohdy špatnému skloňování nebo nedokončení slov.

Úroveň publikační činnosti doktoranda a kvalita použitých informačních zdrojů

Disertant je spoluautorem sedmi publikací ve sbornících vědeckých konferencích a jedné publikaci s přiznaným IF=3.847 (Q2). Disertant čerpal ze 114 kvalitních časopiseckých a internetových informačních zdrojů.

Drobné připomínky:

Str. 33 Autor uvádí na obr. 4h Fraktální marker, ale marker nevykazuje např. soběpodobnost, což je klíčový rys fraktálů.

V DP autor používá termín *návrh design marker*, proč?

Jakou funkci má “bounded box”?

Problematika řešená v DP je specifická, co je Vaše a co jste převzal?

Závěr:

Disertační práce **Ing. Milana Košťáka** je napsaná přehledně, logicky je členěna do kapitol, které na sebe významově navazují, přináší nová řešení v oblasti rozšířené reality, plně zapadá do problematiky studijního oboru a proto doporučuji disertační práci k obhajobě.

V Pardubicích dne 26. 5. 2023

