

# TOWARDS CLINICAL TRANSLATION OF RAMAN SPECTROSCOPY FOR TUMOR CELL IDENTIFICATION

DISSERTATION

zur Erlangung des akademischen Grades *doctor rerum naturalium* (*Dr. rer. nat.*)



**FRIEDRICH-SCHILLER-  
UNIVERSITÄT  
JENA**

Vorgelegt dem Rat der Chemisch-Geowissenschaftlichen Fakultät der  
Friedrich-Schiller-Universität Jena

von *M. Sc. Roman Kiselev*

Geboren am 17. Juni 1989 in Leningrad

---

Jena    April 2019

---

# Contents

<b>1. Introduction</b>	<b>13</b>
1.1. Circulating Tumor Cells . . . . .	13
1.2. Detection Methods of Circulating Tumor Cells . . . . .	15
1.3. Motivation and Goals of the Work . . . . .	19
<b>2. The Raman Effect</b>	<b>21</b>
2.1. Short Theory of Raman Scattering . . . . .	21
2.2. Raman Spectroscopy for Clinical Diagnostics . . . . .	24
<b>3. Instrumentation</b>	<b>25</b>
3.1. Design Considerations for Raman Microscopes . . . . .	25
3.1.1. Generalized Optical Microscope . . . . .	26
3.1.2. Selection of the Excitation Wavelength for Raman Spectroscopy . . . . .	27
3.1.3. Confocality and Signal Collection Efficiency . . . . .	29
3.1.4. Raman Microscope Alignment . . . . .	32
3.2. Raman Instruments Used in this Work . . . . .	33
3.2.1. Flexible Raman Instrument . . . . .	33
3.2.2. RS660 Instrument . . . . .	34
3.2.3. RS785 Instrument . . . . .	35
3.3. Characterization of Raman Systems . . . . .	37
3.3.1. RS785 instrument . . . . .	37
3.3.2. RS660 Instrument . . . . .	39
3.4. Cell Handling with Microfluidics . . . . .	42
3.4.1. Microhole Array Microfluidic Chip . . . . .	43
3.4.2. “CanDo” Microfluidic Chip “Cartridge I” . . . . .	45
3.4.3. RoC Quartz Microfluidic Chip . . . . .	47
3.5. Discussion of Raman Instrumentation . . . . .	54
<b>4. Integration of Raman Instruments and Data Management</b>	<b>57</b>
4.1. Data Representation in a Spectroscopic Experiment . . . . .	57
4.1.1. Multivariate structure of the spectral data . . . . .	57
4.1.2. Long representation of the spectral data . . . . .	58
4.2. Keeping experimental data organized within a database . . . . .	59
4.2.1. Database Modeling for Storage of Spectral Data . . . . .	59
4.2.2. Constraints and mechanism of transactions . . . . .	60
4.2.3. Network security . . . . .	61

4.3.	Data acquisition software . . . . .	63
4.3.1.	Graphical user interface . . . . .	63
4.3.2.	Save/Recall Cell Locations . . . . .	66
4.3.3.	Calibration of the Laser Power . . . . .	66
4.3.4.	RS785 specific features . . . . .	67
4.3.5.	RS660 specific features . . . . .	68
4.4.	Algorithm for Automatic Wavelength Calibration . . . . .	69
4.4.1.	Basics of Wavelength Calibration for Dispersive Spectrometers . . . . .	69
4.4.2.	Peak Detection Using Morphological Operations . . . . .	70
4.4.3.	Matching of Peaks with Atomic Emission Lines . . . . .	71
4.4.4.	Outlook for Algorithm Improvement . . . . .	74
4.5.	Database access with <code>db2spc</code> . . . . .	74
4.5.1.	Automatic Dark Frame Subtraction . . . . .	75
4.5.2.	Use of cell images stored in the database . . . . .	75
4.5.3.	Interactive Database Viewer . . . . .	80
4.6.	Improved Data Management with the Spectra Database . . . . .	81
4.7.	Overview of the Collected Dataset . . . . .	83
<b>5.</b>	<b>Data analysis</b>	<b>87</b>
5.1.	Detection of Outliers in Raman Spectra . . . . .	87
5.1.1.	New Algorithm for Automatic Detection of Outliers in Spectral Data . . . . .	89
5.1.2.	Calculation of Average Penalty $\bar{S}$ . . . . .	90
5.2.	Automatic Cell Identification with Machine Learning . . . . .	92
5.3.	Unsupervised Methods and their Application to Spectra . . . . .	94
5.4.	Supervised Machine Learning for Cell Identification . . . . .	97
5.4.1.	Validation . . . . .	99
5.5.	Classification Models for Cell Identification . . . . .	101
5.5.1.	Selection of the Optimal Number of Principal Components . . . . .	104
5.5.2.	Accuracy of Classifiers Depending on the Exposure Time . . . . .	105
5.6.	Use of Classification Models to Identify Specific Tumor Type . . . . .	109
5.6.1.	Prediction of Patient Samples . . . . .	112
5.7.	Experimental Demonstration of Cell Identification Workflow . . . . .	114
5.8.	Automation and Data Management Pave the Way Towards Biospectroscopic Clinical Cancer Diagnostics . . . . .	116
	<b>Bibliography</b>	<b>116</b>
	<b>Appendices</b>	<b>123</b>
A.	Quantum efficiency of the used cameras . . . . .	123
B.	Fluorescence Attachment for Raman System . . . . .	124
C.	Full Database Structure . . . . .	125
D.	Development of import filters for SPE file format . . . . .	126
E.	Usage examples of <code>db2spc</code> package . . . . .	128
F.	Examples of images stored in the database . . . . .	131